

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/05621

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 16, 19 and 22 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims: it is covered by claims Nos.:

1-23 (PARTIALLY)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.



INTERNATIONAL SEARCH REPORT

International Application No

PC 00/05621

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/00 C07K14/47 G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	WO 99 22243 A (ENDRESS GREGORY A ;FLORENCE KIMBERLY A (US); HUMAN GENOME SCIENCES) 6 May 1999 (1999-05-06) abstract page 101 -page 102 Seq.Id.No.215 ---	1-23
A	SHIROZU M ET AL: "CHARACTERIZATION OF NOVEL SECRETED AND MEMBRANE PROTEINS ISOLATED BY THE SIGNAL SEQUENCE TRAP METHOD" GENOMICS,US,ACADEMIC PRESS, SAN DIEGO, vol. 37, no. 3, 1 November 1996 (1996-11-01), pages 273-280, XP002054773 ISSN: 0888-7543 the whole document --- -/-	1-23

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

1 September 2000

Date of mailing of the international search report

05 1. 01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040. Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Panzica, G



INTERNATIONAL SEARCH REPORT

International Application No

PC 00/05621

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JACOBS K A ET AL: "A genetic selection for isolating cDNAs encoding secreted proteins" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES,GB,ELSEVIER SCIENCE PUBLISHERS, BARKING, vol. 198, no. 1-2, 1 October 1997 (1997-10-01), pages 289-296, XP004116069 ISSN: 0378-1119 the whole document	1-23
A	--- WO 98 30696 A (GENETICS INST) 16 July 1998 (1998-07-16) the whole document	1-23
A	--- WO 98 32853 A (GENETICS INST) 30 July 1998 (1998-07-30) abstract	1-23
A	--- WO 98 40486 A (GENETICS INST) 17 September 1998 (1998-09-17) abstract -----	1-23



FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-23 (partially)

An isolated polypeptide having sequence as set forth in Seq.Id.No.1 of the sequence listing and its nucleic acid sequence as set forth in Seq.Id.No.23. Methods of production and uses thereof.

2. Claims: 1-23 (partially)

Inventions 2-22:

Same as for invention no.1 but respectively to each following pair of aminoacid sequences with their respective nucleic acid sequences:

Invention 2: Seq.Id.Nos.2 and 24.

Invention 3: Seq.Id.Nos.3 and 25.

....

Invention 22: Seq.Id Nos.22 and 44.



INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PC 00/05621

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9922243	A	06-05-1999	AU 1273499 A	17-05-1999
			EP 1042674 A	11-10-2000
			AU 1118499 A	17-05-1999
			WO 9921575 A	06-05-1999

WO 9830696	A	16-07-1998	US 5945302 A	31-08-1999
			US 5972652 A	26-10-1999
			AU 2723397 A	29-10-1997
			EP 0907732 A	14-04-1999
			EP 1007660 A	14-06-2000
			JP 2000508172 T	04-07-2000
			AU 5822398 A	03-08-1998
			AU 2558599 A	26-07-1999
			EP 1047713 A	02-11-2000
			WO 9935168 A	15-07-1999

WO 9832853	A	30-07-1998	AU 5828398 A	18-08-1998
			EP 1007661 A	14-06-2000

WO 9840486	A	17-09-1998	US 5976837 A	02-11-1999
			AU 6702298 A	29-09-1998
			EP 0973890 A	26-01-2000



INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 92/2376

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.5, 23.1; 435/320.1, 440, 252.3, 69.1, 7.1; 530/350, 387.1; 514/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ADAMS et al. Complementary DNA sequencing: Expressed sequence tags and the human genome project. Science. 21 June 1991, Vol. 252, pages 1651-1656, see entire document.	1-22

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 13 JANUARY 1999	Date of mailing of the international search report 03 FEB 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer <i>D. Lawrence For</i> JAMES MARTINELL Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/22376

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 23
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Claim 23 is directed to a product of the process of claim 20. claim 20 is not a process for the production of a product, but a process for the detection of a substance. hence, no meaningful search can be carried out.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP92/22376

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

G01N 33/68, 33/53; C07K 16/00; C12N 15/11, 15/12, 15/00, 15/63; A61K 38/17, 38/16; C12P 21/02

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

536/23.5, 23.1; 435/320.1, 440, 252.3, 69.1, 7.1; 530/350, 387.1; 514/12

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, STN, MPSRCH (SEQ ID NOs 11 and 160 only). One nucleotide sequence and one amino acid sequence have been searched. It is not clear which sequences are embraced by the claims because the claims refer to sequences X and Y. The table beginning after page 209 contains many sequences X and Y, yet the claims refer to X and Y in the singular. If the claims are to embrace more than one X and more than one Y, it is not clear whether each X always requires the corresponding sequence Y. Additionally, the claims are in improper format in referring to the description (see PCT Rule 6.2(a)). Accordingly, the first X nucleotide sequence disclosed and the first Y amino acid sequence mentioned in the claims were searched.



PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference PF-0675 PCT	<div style="display: flex; justify-content: space-between;"> <div style="text-align: center;">FOR FURTHER ACTION</div> <div style="font-size: small;">see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.</div> </div>	
International application No. PCT/US 00/ 05621	International filing date (day/month/year) 03/03/2000	(Earliest) Priority Date (day/month/year) 05/03/1999
Applicant INCYTE PHARMACEUTICALS, INC. et al.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 5 sheets.
☐ It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

☐ the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

☒ contained in the international application in written form.

☐ filed together with the international application in computer readable form.

☐ furnished subsequently to this Authority in written form.

☒ furnished subsequently to this Authority in computer readable form.

☒ the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

☒ the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. ☒ **Certain claims were found unsearchable** (See Box I).

3. ☒ **Unity of invention is lacking** (see Box II).

4. With regard to the **title**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No.

☐ as suggested by the applicant.

☐ because the applicant failed to suggest a figure.

☐ because this figure better characterizes the invention.

☐ None of the figures.



PATENT COOPERATION TREATY

PCT

Rec'd 02 APR 2002

WIPO

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

8

Applicant's or agent's file reference PF-0675 PCT	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT IPEA 416)	
International application No. PCT/US00/05621	International filing date (day/month/year) 03 March 2000 (03.03.2000)	Priority date (day/month/year) 05 March 1999 (05.03.1999)
International Patent Classification (IPC) or national classification and IPC IPC(7): C12N 1/68; G01N 33/53 and US Cl.: 435.6, 69.1, 320.1; 536.23.1; 530.300; 424.184.1		
Applicant INCYTE PHARMACEUTICALS, INC.		
<p>1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.</p> <p>2. This REPORT consists of a total of <u>5</u> sheets, including this cover sheet.</p> <p><input type="checkbox"/> This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).</p> <p>These annexes consist of a total of <u>0</u> sheets.</p>		
<p>3. This report contains indications relating to the following items:</p> <p>I <input checked="" type="checkbox"/> Basis of the report</p> <p>II <input type="checkbox"/> Priority</p> <p>III <input checked="" type="checkbox"/> Non-establishment of report with regard to novelty, inventive step and industrial applicability</p> <p>IV <input type="checkbox"/> Lack of unity of invention</p> <p>V <input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement</p> <p>VI <input type="checkbox"/> Certain documents cited</p> <p>VII <input type="checkbox"/> Certain defects in the international application</p> <p>VIII <input checked="" type="checkbox"/> Certain observations on the international application</p>		
Date of submission of the demand 26 September 2000 (26.09.2000)	Date of completion of this report 27 February 2002 (27.02.2002)	
Name and mailing address of the IPEA-US Commissioner of Patents and Trademarks Box PCT Washington, DC 20231	Authorized officer <i>James K. Feman</i> Telephone No. 703 308 0196	

Form PCT IPEA 400 (cover sheet) July 1998



I. Basis of the report

1. With regard to the elements of the international application *

☒ the international application as originally filed.☒ the description:

pages 1-74 as originally filed

pages NONE, filed with the demand

pages NONE, filed with the letter of _____

☒ the claims:

pages 75-77, as originally filed

pages NONE, as amended (together with any statement) under Article 19

pages NONE, filed with the demand

pages NONE, filed with the letter of _____

☐ the drawings:

pages NONE, as originally filed

pages NONE, filed with the demand

pages NONE, filed with the letter of _____

☒ the sequence listing part of the description:

pages 1-28 as originally filed

pages NONE, filed with the demand

pages NONE, filed with the letter of _____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language _____ which is:☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).☐ the language of publication of the international application (under Rule 48.3(b)).☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:☒ contained in the international application in printed form.☒ filed together with the international application in computer readable form.☐ furnished subsequently to this Authority in written form.☐ furnished subsequently to this Authority in computer readable form.☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.4. ☐ The amendments have resulted in the cancellation of☐ the description, pages NONE☐ the claims, Nos. NONE☐ the drawings, sheets fig NONE5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)) **

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments. (Rules 70.16 and 70.17)

** Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report



INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US00/0562

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The question whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been and will not be examined in respect of:

- ☐ the entire international application,
- ☒ claims Nos. Claims 1-23 in part, as they read on SEQ ID NO. 2, 22, 24-44

because:

- ☐ the said international application, or the said claim Nos. _____ relate to the following subject matter which does not require international preliminary examination (*specify*):

- ☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. _____ are so unclear that no meaningful opinion could be formed (*specify*):

- ☐ the claims, or said claims Nos. _____ are so inadequately supported by the description that no meaningful opinion could be formed.

- ☒ no international search report has been established for said claims Nos. 1-23 in part, SEQ 2, 22, 24-44

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

- ☐ the written form has not been furnished or does not comply with the standard.
- ☒ the computer readable form has not been furnished or does not comply with the standard.



V. Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**1. STATEMENT**

Novelty (N)	Claims <u>NONE</u>	YES
	Claims <u>1-23</u>	NO
Inventive Step (IS)	Claims <u>NONE</u>	YES
	Claims <u>1-23</u>	NO
Industrial Applicability (IA)	Claims <u>1-23</u>	YES
	Claims <u>NONE</u>	NO

2. CITATIONS AND EXPLANATIONS

The sequences disclosed in the priority document and the related computer readable format are not the same as the sequences in the application, and are not in the same order as the sequences in this application. Therefore, the application is not entitled to the to the priority date.

Documents Cited: WO 99/22243, identified in Search Report.

Claims 1-23 lack novelty under PCT Article 33(2) as being anticipated by WO 99/22243 (HUMAN GENOME SCIENCES) (06 May 1999). This publication discloses SEQ ID NO: 215 having identity to SEQ ID NO: 1 and/or 23. Polypeptides, fragments, and active fragments are disclosed, as are polynucleotides, vectors, transformed host cells. Methods of making and using the polynucleotide and/or polypeptide are disclosed.

Claims 1-23 meet the criteria set out in PCT Article 33(4) for industrial applicability.



VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the questions whether the claims are fully supported by the description, are made:

The description is objected to under PCT Rule 66.2(a)(v) as lacking clarity under PCT Article 5 because it fails to adequately enable practice of the claimed invention because: no biological activity has been ascribed to the polypeptide of SEQ ID NO: 1. In Table 2, only a predicted signal sequence is identified. Table 3 discloses that short fragments of SEQ ID NO: 23 appear to be expressed in musculoskeletal tissue and reproductive tissue, and indicates general categories of potentially related diseases. No specific data regarding the claimed sequences and any particular biological activity is disclosed. Lacking the knowledge of a particular biological activity, one would not be able to make or identify biologically active fragments of the polypeptide of SEQ ID NO: 1. Similarly, the description fails to adequately enable methods of diagnosing and treating a disease. If the specific biological activity, and its relationship to a named disease or medical condition is not described, one of skill in the art would not be able to practice the claimed methods.

Claims 1-23 are objected to as lacking clarity under PCT Rule 66.2(a)(v) because practice of the claimed invention is not enabled as required under PCT Rule 5.1(a) for the reasons set forth in the immediately preceding paragraph.



PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Commissioner
 US Department of Commerce
 United States Patent and Trademark
 Office, PCT
 2011 South Clark Place Room
 CP2/5C24
 Arlington, VA 22202
 ETATS-UNIS D'AMERIQUE
 in its capacity as elected Office

Date of mailing (day/month/year) 04 May 2001 (04.05.01)	
International application No. PCT/US00/05621	Applicant's or agent's file reference PF-0675 PCT
International filing date (day/month/year) 03 March 2000 (03.03.00)	Priority date (day/month/year) 05 March 1999 (05.03.99)
Applicant TANG, Y., Tom et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:
26 September 2000 (26.09.00)

☐ in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was
☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO
 34, chemin des Colombettes
 1211 Geneva 20, Switzerland

Authorized officer

Claudio Borton

Facsimile No.: (41-22) 740 14 35

Telephone No.: (41-22) 348 83 38

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

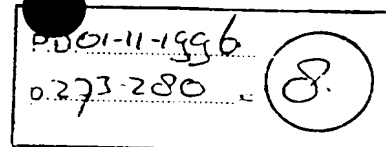
196

197

198

199

200



XP-002054773

Characterization of Novel Secreted and Membrane Proteins Isolated by the Signal Sequence Trap Method

MICHIO SHIROZU,¹ HIDEAKI TADA,¹ KEI TASHIRO, TOMOYUKI NAKAMURA, NELSON D. LOPEZ, MARTINA NAZAREA, TSUNEYOSHI HAMADA, TOSHIHIKO SATO, TORU NAKANO,² AND TASUKU HONJO³

Department of Medical Chemistry, Kyoto University Faculty of Medicine, Sakyo-ku, Kyoto 606 Japan

Received March 11, 1996; accepted July 9, 1996

We recently described a method, called the signal sequence trap (SST) method, to clone cDNAs of secreted proteins and/or type I transmembrane proteins containing N-terminal signal sequences by using an epitope-tagging expression plasmid vector. In this paper we describe the summary of a large-scale screening of approximately 5900 clones of an SST cDNA library constructed from mouse bone marrow stromal cell line ST-2 cells. Of 26 positive clones obtained and sequenced, 11 clones appeared to contain authentic signal sequences. Five of the clones corresponded to the 5' ends of the cDNA of known genes containing N-terminal signal sequences. The full-length cDNA clones of the 6 other unknown clones were isolated and sequenced. One clone, termed SDF3, encoded a mouse counterpart of human pigment epithelium-derived factor. Another clone, termed SDR1, had considerable homology with basigin, a member of the immunoglobulin superfamily. A third clone, termed SDF5, had partial homology with a *Drosophila* tissue polarity gene *frizzled* (*fz*) and its rat homologues, *fz-1* and *fz-2*. The other three clones had no significant homology with sequences in the databases. These results indicate that the SST method is effective and useful for the isolation of secreted and membrane proteins without knowledge of their functions. © 1996 Academic Press, Inc.

INTRODUCTION

The elucidation of molecular mechanisms for intercellular signaling and cell adhesion is essential for un-

derstanding the mechanism for the development and differentiation of multicellular organisms. Although a large number of molecules involved in signaling or adhesion have been cloned, there still remain many unknown molecules that are important for these functions. Most of the molecules involved in signaling or adhesion are secreted or membrane-anchored proteins, such as growth factors, hormones, neuropeptides, and their receptors or adhesion molecules. Many of these proteins contain hydrophobic signal sequences at the N-termini of the precursor forms. To establish a general cDNA cloning strategy for these secreted or membrane-anchored proteins, we developed a new cloning method for the isolation of cDNA fragments encoding N-terminal signal sequences (Tashiro *et al.*, 1993). cDNA fragments encoding signal sequences were detected by surface expression of the IL-2 receptor α chain (Tac) (Uchiyama *et al.*, 1981) as fusion proteins in transfected cells. Using this method, termed the signal sequence trap (SST) method, we have so far reported the isolation of cDNA for a member of the α -chemokines, SDF1 α/β , and a growth factor receptor lymphotoxin β receptor, by the SST method (Tashiro *et al.*, 1993; Nakamura *et al.*, 1995), suggesting that this method may be widely useful.

The hematopoietic system is one of the best characterized organs for analyzing the mechanisms of mammalian cell growth and differentiation. Bone marrow stromal cell lines support hematopoiesis from bone marrow stem cells by secretion of cytokines and by interaction with cell-adhesion molecules. Many molecules involved in cell-cell interactions during lymphohematopoiesis have been cloned and are known to be expressed in bone marrow stromal cell lines (Metcalf, 1989; Paul *et al.*, 1991). However, there are still a number of unknown surface molecules involved in the early phases of hematopoiesis. We therefore performed a large-scale screening of cDNAs encoding secreted and transmembrane proteins from a bone marrow stromal cell line ST-2 (Nishikawa *et al.*, 1988) using the SST method. Isolated cDNAs will be expressed as proteins and their functions could be tested in an inductive system of lymphohematopoiesis from embryonic stem cells (Nakano *et al.*, 1994).

The nucleotide sequences reported in this paper have been deposited with the GenBank/EMBL Data Libraries under Accession Nos. SDF3 (Sdf3), D50460; SDF4 (Sdf4), D50461; SDF5 (Sdf5), D50462; SDR1 (Sdr1), D50463; and SDR2 (Sdr2), D50464.

¹ These authors contributed equally to this work.

² Present address: Department of Molecular Cell Biology, Research Institute for Microbial Diseases, Osaka University, Yamada-Oka, Suita, Osaka 565 Japan.

³ To whom correspondence should be addressed at Department of Medical Chemistry, Kyoto University Faculty of Medicine, Yoshida, Sakyo-ku, Kyoto 606, Japan. Telephone: 81-75-753-4371. Fax: 81-75-753-4388.

TABLE 1
Summary of the Trapped cDNA Clones of SST

Library	Clone	ORF in frame to Δ TAC	Homology		Hydrophobic peak	Stop codon upstream of initiation codon	Located at the 5' end of full-length cDNA	Signal sequence ^a	Name of full-length clone
			Known sequences	N-terminal SS					
Library I	H1	+	+	+	Short			T	Osteonectin
	H2	-						F	
	H3	+	-					U	
	H4	+	+	-				F	Peptidyl-glycine α -amidating monooxygenase
	S1	+	-		+	-	+	T	SDR1
	S2	+	-		+	-	+	T	SDF3
	S3	+	-		+			U	
	S4	+	-		-			F	
	S6a	+	-		+	+	+	T	SDR2
	T1	+	-		+	-	+	T	SDF2
	T2	+	-		Short			U	
	T4	+	-		-			F	
	T5	+	+	+				T	Ribophorin 1
	T7	+	-					F	
	T8	+	-		Short			U	
	E6a	+	+	+				T	Protein disulfide isomerase
Library II	F7a	+	-		+	-	-	F	
	NT1	+	-		+			U	
	NT2	+	+	+				T	Fibronectin
	NT3	+	-		+	+	+	T	SDF4
	NT4	+	-		-			F	
	NT5	+	-		+	+	+	T	SDF5
	NT6	-						F	
	P5c	-						F	
	R5g	+	-		+	+	-	F	
	TA	+	+	+	+			T	Osteonectin

^a T, true; F, false; U, undecided.

We will summarize the results of the large-scale screening of the SST library of ST-2 cells to examine the ratio of clones with the authentic signal sequence to artifacts and to test if this strategy is valid as a general cloning method for isolation of signal sequence-containing molecules. Ten clones with authentic signal sequences were obtained from approximately 5900 cDNA clones. The six full-length cDNA clones encoding novel proteins were isolated, sequenced, and characterized.

MATERIALS AND METHODS

Construction and screening of the SST library. Construction of the SST library was performed as described previously (Tashiro *et al.*, 1993) except that two sets of sonicated cDNA fragments with two size ranges (one between 300 and 500 bp and the other between 500 and 700 bp) were selected and used as inserts of the SST library. About 10,000 clones of the shorter insert cDNA library (library I) were plated on LB agar plates, and colony hybridization was performed with the probes of GlA (Ikeda *et al.*, 1991), JE (Rollins *et al.*, 1988), and SDF1 according to the conventional method (Sambrook *et al.*, 1989). Colonies that did not hybridize with any of the probes were picked up and spotted on 9-cm agar plates with a matrix format (7 rows by 7 lines) (Tashiro *et al.*, in press). Approximately 3500 preselected colonies were screened after 49 clones were pooled as

described before (Tashiro *et al.*, 1993). The longer insert library (library II) was screened without preselection.

DNA sequencing. DNA sequencing was carried out using double-stranded templates by the dideoxy method using the *Taq* DyeDeoxy terminator cycle sequencing kit and an automated DNA sequencer (373A, Applied Biosystems). The sequence information was used for analysis of hydrophobicity and compared with sequences in the GenBank database for homology using BLAST and FASTA.

RNA extraction and Northern blot analysis. Crude RNAs of mouse tissues were extracted by using an RNA extraction kit (TRIzol, GIBCO BRL), and poly(A)⁺ RNAs were enriched by oligo(dT) latex (Oligotex(dT) Super, Takara). Two micrograms of mouse tissue poly(A)⁺ RNA was electrophoresed through a 1.0% agarose gel and transferred to a nylon membrane (BIODYNE, Pall) in 10 \times SSC (1 \times SSC is 0.15 M NaCl, 15 mM sodium citrate) overnight, and cross-linked with a UV crosslinker (Funalinker, Funakoshi). The membrane was serially used for hybridization with the SDF3, SDF4, SDF5, SDR1, SDR2, and mouse β -actin probes radiolabeled by the random priming method (Sambrook *et al.*, 1989). The probes for the new genes were full-length cDNAs. Hybridization was performed according to a conventional method, and washing was performed under stringent conditions (Sambrook *et al.*, 1989). Autoradiograms were analyzed with a Bio-image analyzer (BAS 2000, Fuji Film).

RESULTS

cDNA Library Screening

In the previous study, a small-scale screening (approximately 600 clones) of the SST library with 300- to

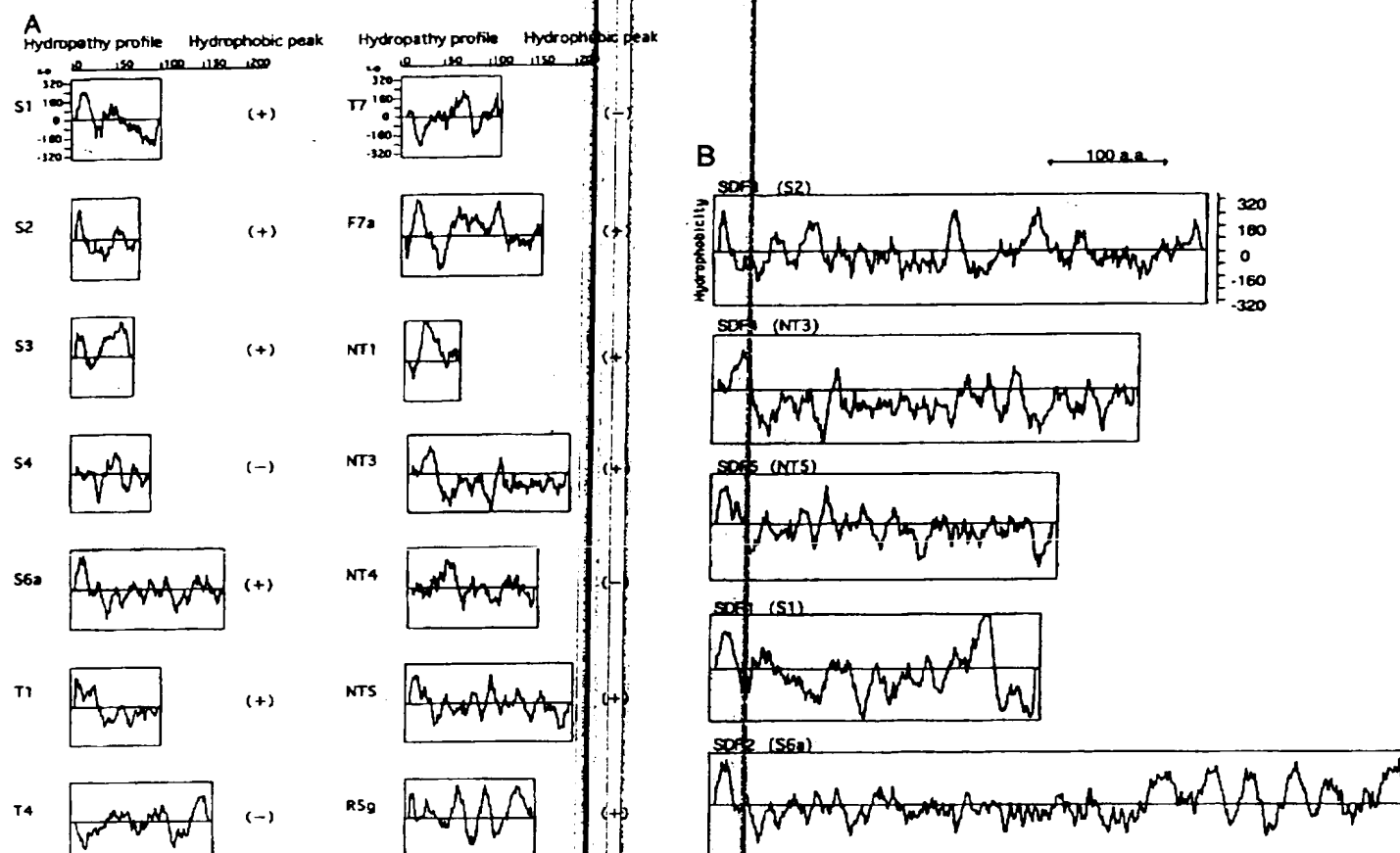


FIG. 1. (A) Hydropathy profiles of the signal sequence trapped clones. Hydropathy profiles of the 14 trapped clones excluding those identical to known genes or carrying short ORFs (fewer than 35 amino acid residues) are shown. (+) or (-) indicates a clone with or without a hydrophobic peak. (B) Hydropathy profiles of the full-length cDNA of newly identified clones. Hydropathy profiles were analyzed with a computer program based on Kyte and Doolittle (1982).

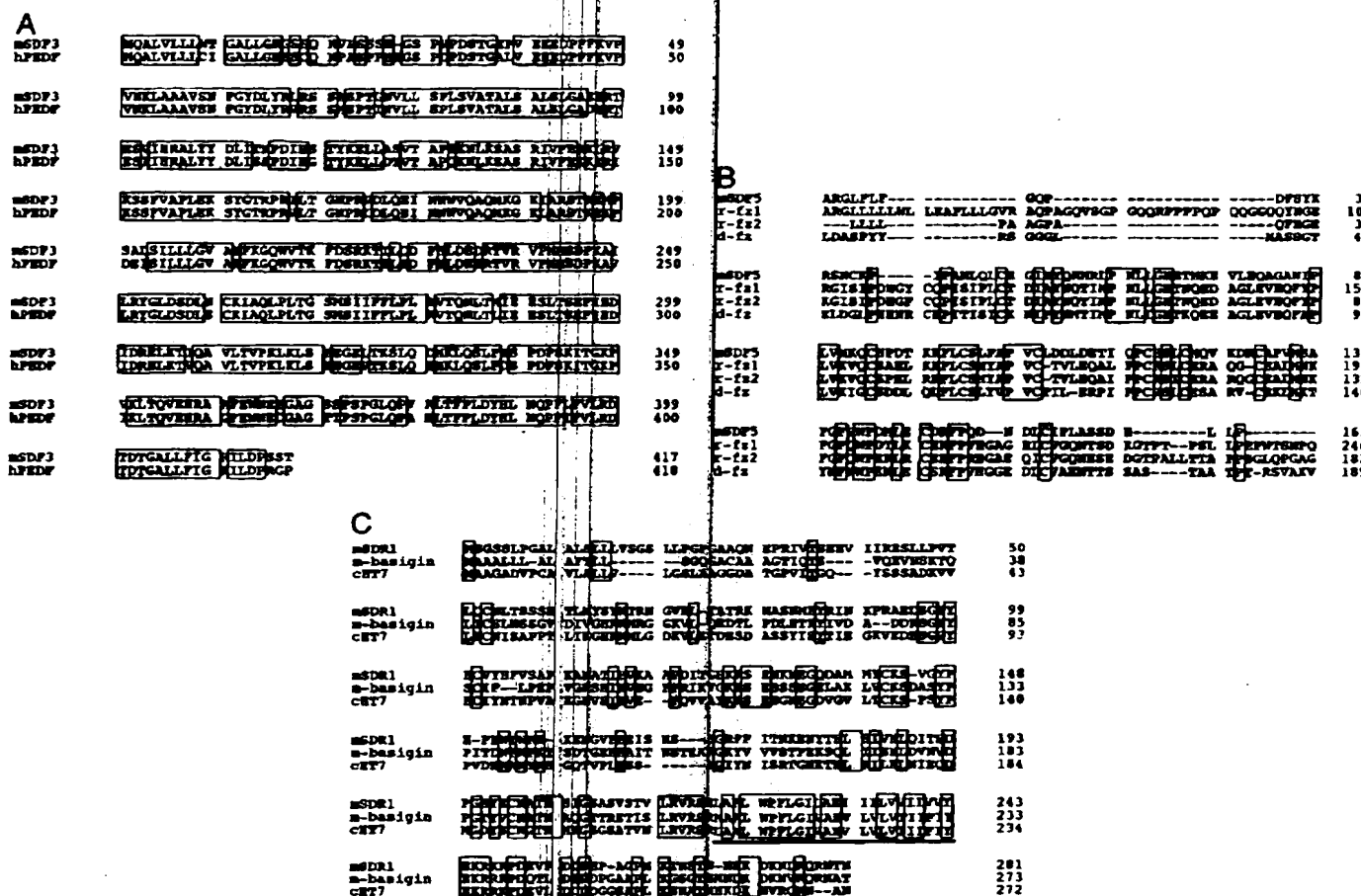
500-bp inserts (Tashiro *et al.*, 1993) led to the repeated isolation of three clones, Gla, JE, and SDF1. To exclude such abundant clones two strategies were introduced: (a) elimination of the abundant clones by colony hybridization (library I) and (b) construction of an SST library

with longer inserts (library II), which should remove the three clones mentioned above because their coding sequences are less than 500 bp. The two strategies are complementary to some extent, because the former is tedious but does not exclude other shorter cDNA clones

TABLE 2

Summary of Characterization of the cDNA Clones Obtained by ST-2 SST

Clone name	cDNA length (kb)	ORF length (aa)	Homology	Sizes of mRNA (kb)	Tissue distribution of mRNA	References
SDF1	3.4	93	α -Chemokine	3.4, 1.8	Ubiquitous (peripheral blood leukocyte-)	Tashiro <i>et al.</i> , 1993
SDF2	1.0	211	Yeast Pmt1p and Pmt2p	1.0	Ubiquitous	Hamada <i>et al.</i> , in press
SDF3	1.4	417	Pigment epithelium-derived factor	1.6	Liver++, brain+, heart+, muscle+, spleen+, thymus+	This paper
SDF4	1.8	356	4 EF hands	5.0, 1.8	Ubiquitous	This paper
SDF5	1.8	295	<i>Drosophila</i> frizzled rat frizzled-1 and -2	4.0, 2.2, 1.5	Heart+, kidney+, lung+, brain±, thymus±	This paper
SDR1	1.9	281	Ig superfamily basigin chicken BT-7	2.6, 2.1	2.1 kb, Ubiquitous	This paper
SDR2	2.4	592		6.8, 2.6	2.6 kb, brain-specific Kidney+, liver±, spleen±,	This paper



while the latter might miss clones with open reading frames (ORFs) shorter than 500 bp.

About 3500 clones of library I were screened as described before (Tashiro *et al.*, 1993) and 15 positive clones were obtained. On the other hand, 11 positive clones were obtained by screening about 2400 clones of library II. The ratios of positive clones to artifact clones were similar between the two libraries and all of the 26 clones isolated will be characterized below.

Sequence Analysis of Positive SST cDNA Clones

The nucleotide and amino acid sequences of the inserts of the 26 positive clones were determined and compared with sequences in the GenBank database for homology. As summarized in Table 1, there were 6 clones identical to partial sequences of known genes as follows: H1 and TA, osteonectin precursor (Howe *et al.*, 1988); H4, peptidyl-glycine α -amidating monooxygenase precursor (Eipper *et al.*, 1992); T5, ribophorin 1 precursor (rough endoplasmic reticulum protein) (Behal *et al.*, 1990); E6a, protein disulfide isomerase pre-

cursor (rough endoplasmic reticulum protein) (Gong *et al.*, 1988); NT2, fibronectin precursor (Schwarzbauer *et al.*, 1987). Two signal sequence-trapped clones, H1 and TA, corresponded to the 5' portions of osteonectin precursor cDNA, but the inserts of the two clones were not identical because they were derived from two distinct libraries. The hydropathy profiles of 14 clones are shown in Fig. 1A, whereas those of the other 12 clones are identical to known genes or carrying no or very short ORFs (fewer than 35 amino acid residues) are excluded

Structural examination of the 26 trapped cDNA clones did not allow us to conclude that the following 15 clones have authentic signal sequences. Three clones, H2, NT6, and P5c, had no ORFs in frame with the Tac cDNA sequence of the cloning vector. H4 was identical to part of a known gene, but it did not correspond to the N-terminal signal sequence (data not shown). S4, T4, T7, and NT4 had no obvious hydrophobic peaks following the putative first methionine residues. H3, T2, and T8 had very short ORFs (fewer than 35 aa) though their sequences had hydrophobic proper

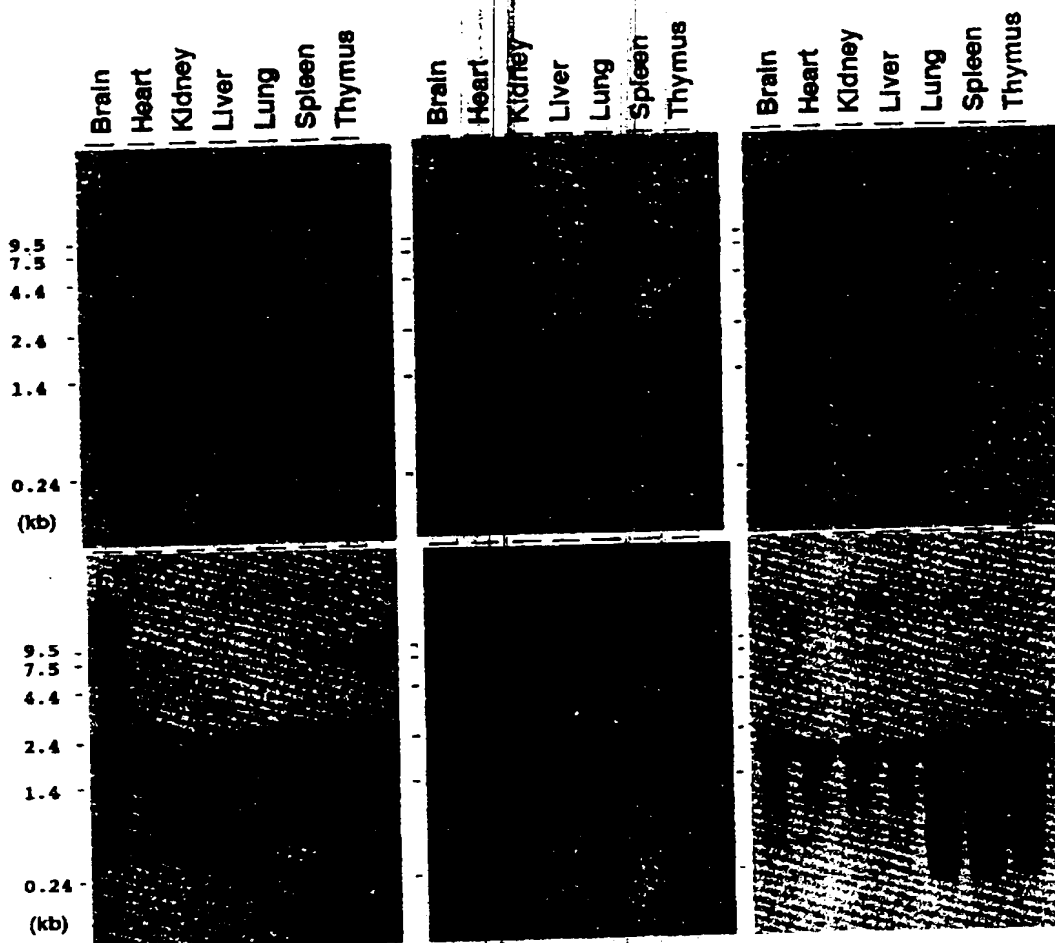


FIG. 3. Tissue distribution of SDF3, SDF4, SDF5, SDR1, and SDR2 mRNAs. The same Northern blot filter containing 2 μ g of poly(A)⁺ RNA from heart, brain, kidney, liver, lung, spleen, and thymus was hybridized serially with a DNA probe of SDF3 (A), SDF4 (B), SDF5 (C), SDR1 (D), SDR2 (E), or β -actin (F) cDNA. The filter was washed after each exposure. Size markers are indicated in kilobases.

ties (data not shown), and therefore it was very difficult to judge whether the sequences were authentic signal sequences or not. As S3 and NT1 had hydrophobic regions far from the first putative methionine residue (72 and 64 residues, respectively), we did not characterize them further although these hydrophobic sequences may serve as signal sequences. The trapped fragments of F7a and R5g did not locate at the 5' ends of the full-length cDNA clones and these clones were not characterized further.

The following five known clones and six unknown clones were concluded to contain authentic signal sequences. Five clones, H1, T5, E6a, NT2, and TA, were identical to the N-terminal signal sequences of known genes (Table 1). The trapped fragments of six clones (S1, S2, S6a, T1, NT3, and NT5) were located at the 5' ends of the full-length cDNA clones.

Sequence Analysis of Full-Length cDNAs of the Newly Identified Clones

Full-length cDNA of the six unknown clones with authentic signal sequences were isolated, sequenced,

and characterized (Table 2). Detailed data of isolation and characterization of the SDF-2 clone were described elsewhere (Hamada *et al.*, in press). The hydropathy profiles of the deduced amino acid sequences were plotted (Fig. 1B). All the clones have the typical N-terminal signal sequences and none of them have KDEL signals for retention in the endoplasmic reticulum (Pelham *et al.*, 1989). A stromal cell-derived receptor (SDR) or stromal cell-derived factor (SDF) was named depending on the presence or absence, respectively, of a putative transmembrane region. SDR1 bears a clear hydrophobic region near the C-terminus and SDR2 has six hydrophobic domains at the C-terminus.

The sequence homology search was performed using GenBank and SWISS-PROT databases. SDF2 has significant homology to putative enzymic regions of yeast dolichyl phosphate-D-mannose: protein mannosyltransferase, Pmt1p, and Pmt2p (Hamada *et al.*, in press). SDF3 has remarkable homology to the human pigment epithelium-derived factor (PEDF) (Steele *et al.*, 1992) (Fig. 2A). There were 78 and 85% identities at the nucleotide and amino acid sequence levels, re-

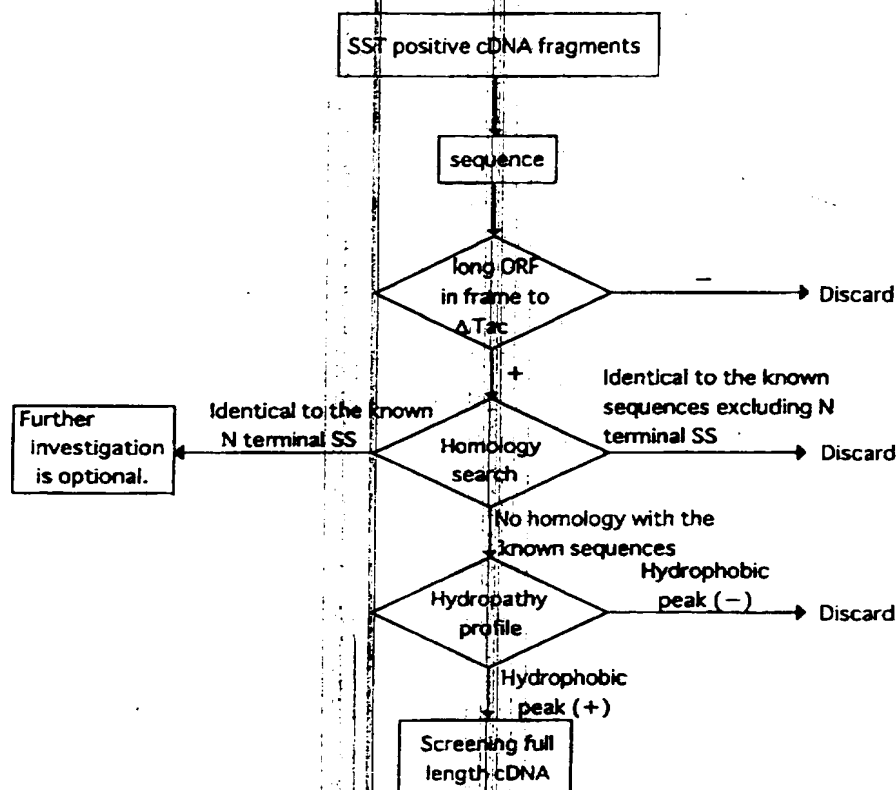


FIG. 4. The strategy for selection of SST positive cDNA fragments. SS, signal sequence.

spectively. This overall similarity strongly suggests that SDF3 is a mouse counterpart for the human PEDF. SDF4 has four EF hand motifs for calcium-binding activity (Kretsinger and Nockolds, 1973). SDF5 shows some homology to the extracellular domains of the *Drosophila* tissue polarity gene *frizzled* (Vinson *et al.*, 1989) and its rat homologues *fz-1* and *fz-2* (Chan *et al.*, 1992) (Fig. 2B), but its C-terminus has no homology with *frizzled*. SDR1 has a considerable homology to mouse basigin (Altruda *et al.*, 1989) and chicken HT7, members of the immunoglobulin superfamily (Fig. 2C). At the amino acid sequence level there is 38% identity between SDR1 and basigin, 38% identity between SDR1 and HT7, and 28% identity among the three proteins, while there is 45% identity between basigin and HT7. SDR2 shows no significant similarity to known genes.

Expression of mRNA of Novel Signal Sequence Trapped Clones

The tissue distribution of SDF3, SDF4, SDF5, SDR1, and SDR2 mRNAs was examined by Northern blot analysis (Fig. 3). SDF3, SDF4, SDR1, and SDR2 mRNAs were expressed in almost all the organs tested, although the expression levels were variable. Several observations are worth noting. SDF3 mRNA was strongly expressed in liver. SDF5 mRNA was expressed in brain, heart, kidney, lung, and thymus, but not in

liver and spleen. A strong 2.1-kb transcript of the SDR1 gene was expressed ubiquitously but a 2.6-kb transcript was detected only in brain. SDR2 mRNA was relatively abundant in kidney.

DISCUSSION

In this report, we described a large-scale screening of two ST-2 SST libraries. Twenty-six positive clones were obtained from 5900 clones from the two libraries. Sequence analysis of the 26 clones revealed that 11 clones contained authentic signal sequences (true), 10 clones were classified as artifacts (false), and 5 clones were classified as undecided (undecided) (Table 1). Artifacts described here are either technical or strategic. The former sequences do not contain signal sequences at all, as judged by the absence of ORFs (H2, NT6, and P5c) or hydrophobic regions (H4, S4, T4, T7, and NT4). The latter sequences do not contain authentic N-terminal signal sequences but do have hydrophobic sequences in the middle of the full-length cDNA clones (F7a and R5g). Undecided clones contain atypical hydrophobic sequences: very short hydrophobic ORFs (H3, T2, and T8) and very long hydrophobic sequences (S3 and NT1). Given the frequencies of artifacts (38%) and novel genes (23%), we conclude that SST is an efficient method for isolating novel secreted or membrane-bound proteins.

After positive clones are screened by the SST method, further discrimination should be required to obtain novel cDNA clones encoding authentic N-terminal signal sequences (Fig. 4). The following strategy is empirically recommended. The first step is sequencing the trapped cDNA clones. It should be verified if the clones have a reasonably long (more than 35 residues) ORF in frame with the Tac sequence of the vector. If not, the clones should be abandoned as artifacts. The next step is a homology search using available databases of published genes and proteins. The clones identical to the known sequences should not be characterized further. The third test is selection by the hydropathy profile. Clones that have no hydrophobic peak at the N-terminus should be eliminated. In this way the clones that appear to contain authentic signal sequences are selected, and their full-length clones are isolated from a full-length cDNA library. The presence of an in-frame stop codon upstream of the first ATG codon is not mandatory because three unknown clones with authentic signal sequences (SDR1, SDF2, and SDF3) did not have upstream in-frame stop codons (see Table 1).

The biggest problem in a random screening approach is to determine the functions of isolated clones. Several strategies will facilitate the understanding of the functions of these molecules: patterns of tissue distribution, overexpression in appropriate cell lines, generation of antibodies against the expressed proteins, and gene targeting. Among the novel genes identified here, we gained some hints about the functions of SDF3 and SDR1 from their homology with known genes. SDF3 had remarkable homology to PEDF, suggesting that SDF3 is the mouse counterpart for human PEDF (Fig. 2A). PEDF is an extracellular neurotrophic agent and a member of the serine protease inhibitor family according to amino acid sequence similarity, but recombinant PEDF did not inhibit any known serine proteases (Becerra *et al.*, 1993). The tissue distribution of human PEDF mRNA had not yet been reported except for RPE cells (Tombran-Tink *et al.*, 1994), but mouse SDF3 mRNA was detected in all the organs tested, suggesting that SDF3 may have other functions in addition to neurotrophic activity. SDR1 belongs to the immunoglobulin superfamily and has considerable homology to mouse basigin and chicken HT7. The mouse basigin was first identified as a membrane protein (gp42) from a mouse fibroblast expression library by screening with a polyclonal antiserum (Altruda *et al.*, 1989). HT7 was identified as an antigen intensely expressed in the blood-brain barrier of the chicken and was closely related to basigin according to the amino acid similarity. As already indicated for basigin and HT7 (Miyauchi *et al.*, 1991), the transmembrane and cytoplasmic domains were highly conserved among the three molecules (Fig. 2C). The amino acid sequences of basigin and HT7 showed 63 and 49% identity, respectively, with that of SDR1. The functions of these molecules are not yet known. The brain-specific 2.6-kb transcript

may be a splicing variant or a product of another related gene.

ACKNOWLEDGMENTS

This work was supported by grants from the Ministry of Education, Science, and Culture of Japan. We are grateful to Dr. K. Shibahara for gifts of mouse tissue poly(A)⁺ RNAs. We also thank Ms. H. Ohori, Ms. S. Okazaki, and Ms. E. Tadokoro for technical and secretarial assistance.

REFERENCES

- Altruda, F., Cervella, P., Gaeta, M. L., Daniele, A., Giaccotti, F., Tanone, G., Stefanuto, G., and Silengo, L. (1989). Cloning of cDNA for a novel mouse membrane glycoprotein (gp42): Shared identity to histocompatibility antigens, immunoglobulines and neural-cell adhesion molecules. *Gene* 85: 445-452.
- Becerra, S. P., Palmer, I., Kumar, A., Steele, F., Shiloach, J., Notario, V., and Chader, G. J. (1993). Overexpression of fetal human pigment epithelium-derived factor in *Escherichia coli*. *J. Biol. Chem.* 268: 23148-23156.
- Behal, A., Prakash, K., D'Eustachio, P., Adesnik, M., Sabatini, D. D., and Kreibich, G. (1990). Structure and chromosomal location of the rat ribophorin I gene. *J. Biol. Chem.* 265: 8252-8258.
- Chao, S. D. H., Karpf, D. B., Fowlkes, M. E., Hooks, M., Bradley, M. S., Vuong, V., Bambino, T., Liu, M. Y. C., Arnaud, C. D., Strewler, G. J., and Nissenson, R. A. (1992). Two homologs of the *Drosophila* polarity gene frizzled (fz) are widely expressed in mammalian tissues. *J. Biol. Chem.* 267: 25202-25207.
- Eipper, B. A., Green, C. B. R., Campbell, T. A., Stoffers, D. A., Keutmann, H. T., Mains, R. E., and Ouafik, L. (1992). Alternative splicing and endoproteolytic processing generate tissue-specific forms of peptidylglycine alpha-amidating monooxygenase (PAM). *J. Biol. Chem.* 267: 4008-4015.
- Gong, Q.-H., Fukuda, T., Parkison, C., and Cheng, S.-Y. (1988). Nucleotide sequence of a full-length cDNA clone encoding a mouse cellular thyroid hormone binding protein (p55) that is homologous to protein disulfide isomerase and the b-subunit of prollyl-4-hydroxylase. *Nucleic Acids Res.* 16: 1203.
- Hamada, T., Tashiro, K., Tada, H., Inazawa, J., Shirozu, M., Shibahara, K., Nakamura, T., Nazarea, M., Nakano, T., and Honjo, T. (1996). Isolation and characterization of a novel secretory protein, stromal cell-derived factor-2 (SDF-2) using the signal sequence trap method. *Gene*, in press.
- Howe, C. C., Overton, G. C., Sawicki, J., Solter, D., Stein, P., and Strickland, S. (1988). Expression of SPARC/osteonectin transcript in murine embryos and gonads. *Differentiation* 37: 20-25.
- Ikedo, T., Yamaguchi, A., Icho, T., Tsuchida, N., and Yoshiki, S. (1991). cDNA and deduced amino acid sequence of mouse matrix bla protein: One of five glutamic acid residues potentially modified to gla is not conserved in the mouse sequence. *J. Bone Miner. Res.* 6: 1013-1017.
- Kretzinger, R. H., and Nockolds, C. E. (1973). Carp muscle calcium-binding protein. II Structure determination and general description. *J. Biol. Chem.* 248: 3313-3326.
- Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157: 105-132.
- Metcalf, D. (1989). The molecular control of cell division, differentiation commitment and maturation in haemopoietic cells. *Nature* 339: 27-30.
- Miyauchi, T., Masuzawa, Y., and Muramatsu, T. (1991). The Basigin group of the immunoglobuline superfamily: Complete conservation of a segment in and around transmembrane domains of human and mouse Basigin and chicken HT7 antigen. *J. Biochem.* 110: 770-774.
- Nakamura, T., Tashiro, K., Nazarea, M., Nakano, T., Sasayama, S.,

- and Honjo, T. (1995). The murine lymphotoxin- β receptor cDNA: Isolation by the signal sequence trap and chromosomal mapping. *Genomics* 30: 312-319.
- Nakano, T., Kodama, H., and Honjo, T. (1994). *Science* 266: 1098-1101.
- Nishikawa, S.-I., Ogawa, M., Nishikawa, S., Kunisada, T., and Kodama, H. (1988). B lymphopoiesis on stromal cell clone: Stromal cell clones acting on different stages of B cell differentiation. *Eur. J. Immunol.* 18: 1767-1771.
- Paul, S. R., Yang, Y. C., Donahue, R. E., Goldring, S., and Williams, D. A. (1991). Stromal cell-associated hematopoiesis: immortalization and characterization of a primate bone marrow-derived stromal cell line. *Blood* 77: 1723-1733.
- Pelham, H. R. B. (1989). Control of protein exit from the endoplasmic reticulum. *Annu. Rev. Cell Biol.* 5: 1-23.
- Rollins, B. J., Morrison, E. D., and Stiles, C. D. (1988). Cloning and expression of JE, a gene inducible by platelet-derived growth factor and whose product has cytokine-like properties. *Proc. Natl. Acad. Sci. USA* 85: 3738-3742.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schwarzbauer, J. E., Patel, R. S., Fonda, D., and Hynes, R. O. (1987). Multiple sites of alternative splicing of the rat fibronectin gene transcript. *EMBO J.* 6: 2573-2580.
- Steele, F. R., Chader, G. J., Johnson, L. V., and Tombran-Tink, J. (1992). Pigment epithelium-derived factor: Neurotrophic activity and identification as a member of the serine protease inhibitor gene family. *Proc. Natl. Acad. Sci. USA* 89: 1526-1530.
- Tashiro, K., Nakano, T., and Honjo, T. (1996). "Methods in Molecular Biology," Humana Press, Totowa, NJ. In press.
- Tashiro, K., Tada, H., Heilker, R., Shirozu, M., Nakano, T., and Honjo, T. (1993). Signal sequence trap: A cloning strategy for secreted proteins and type I membrane proteins. *Science* 261: 600-603.
- Tombran-Tink, J., Pawar, H., Swaroop, A., Rodriguez, I., and Chader, G. J. (1994). Localization of the gene for pigment epithelium-derived factor (PEDF) to chromosome 17p13.1 and expression in cultured human retinoblastoma cells. *Genomics* 19: 266-272.
- Uchiyama, T., Broder, S., and Waldman, T. A. (1981). A monoclonal antibody (anti-Tac) reactive with activated and functionally mature human T cells. I. Production of anti-Tac monoclonal antibody and distribution of Tac (+) cells. *J. Immunol.* 126: 1393-1397.
- Vinson, C. R., Conover, S., and Adler, P. N. (1989). A *Drosophila* tissue polarity locus encodes a protein containing seven potential transmembrane domains. *Nature* 338: 263-264.

A genetic selection for isolating cDNAs encoding secreted proteins

Kenneth A. Jacobs *, Lisa A. Collins-Racie, Maureen Colbert, McKeough Duckett, Margaret Golden-Fleet, Kerry Kelleher, Ronald Kriz, Edward R. LaVallie, David Merberg, Vikki Spaulding, Jen Stover, Mark J. Williamson, John M. McCoy

Genetics Institute, Inc., 87 Cambridge Park Drive, Cambridge, MA 02140, USA

Received 3 March 1997; accepted 14 May 1997; Received by D. Schlessinger

Abstract

We describe a simple, rapid technique for simultaneously isolating large numbers of cDNAs encoding secreted proteins. The technique makes use of a facile genetic selection performed in a strain of *Saccharomyces cerevisiae* deleted for its endogenous invertase gene. A cDNA cloning vector which carries a modified invertase gene lacking its leader sequence is used in conjunction with this strain. Heterologous secreted genes fused appropriately upstream of this defective invertase provide the necessary signals to restore secretion, allowing the yeast to grow on sugars such as sucrose or raffinose. This microbial growth selection facilitates scanning cDNA libraries containing millions of clones, enabling the wholesale identification of novel secreted proteins without the need for specific bioassays. The technique is similar to one previously described (Klein et al. (1996) *Proc. Natl. Acad. Sci. USA* 93, 7108–7113). We describe results using a cDNA library derived from activated human peripheral blood mononuclear cells (PBMC). Genes identified from this library encoded signal sequences of proteins of diverse structure, function, and cellular location such as cytokines, type 1 and type 2 transmembrane proteins, and proteins found in intracellular organelles. In addition, a number of novel secreted proteins were identified, including a chemokine and a novel G-protein-coupled receptor. Since signal sequences possess features conserved throughout evolution, the procedure can be used to isolate genes encoding secreted proteins from both eukaryotes and prokaryotes. © 1997 Elsevier Science B.V.

Keywords: Protein secretion; Invertase; Signal sequence; Chemokine; G-protein-coupled receptor

1. Introduction

Many biological events revolve around intercellular signalling processes mediated by hormones or growth factors and their membrane-bound cellular receptors. Isolation of the genes encoding these proteins is of particular interest, not only for the purpose of understanding the powerful biological processes which these molecules govern, but also for the therapeutic potential which they may hold. Because the repertoire of assays which can be employed for cloning purposes remains incomplete, however, many proteins which regulate

events occurring *in vivo* are beyond the scope of conventional cloning methods. Growth factors and their receptors, however, are all secreted proteins, which suggests an alternative cloning approach.

According to the signal sequence hypothesis (Blobel and Dobberstein, 1975), certain peptide sequences, usually located at the amino-terminus of nascent proteins, determine whether a particular protein is destined for secretion. Experiments confirming this hypothesis have shown that although these signal sequences are unique, they are largely interchangeable among secreted proteins and even between diverse organisms (Hitzeman et al., 1990; Walter and Johnson, 1994; Rapoport et al., 1996). Computational analyses reveal that eukaryotic and prokaryotic signal sequences are remarkably similar, providing a rationale for these observations (von Heijne, 1985). In addition, biochemical experiments have demonstrated considerable functional interchangeability between some components of the eukaryotic and prokaryotic protein secretion pathways (Bernstein et al., 1993; Hartmann et al., 1994).

* Corresponding author. Tel. +1 617 4988105; Fax +1 617 4988878, e-mail: kjacobs@genetics.com

Abbreviations: PBMC, peripheral blood mononuclear cells; TRP, tryptophan; cDNA, DNA complementary to RNA; IFN, interferon; IL, interleukin; GM-CSF, granulocyte monocyte colony stimulating factor; CMD-W, complete minimal plates lacking tryptophan; YPR, yeast extract, peptone, raffinose, antimycin A; oligo, oligodeoxynucleotide.

The Yeast Signal Sequence Trap Vector

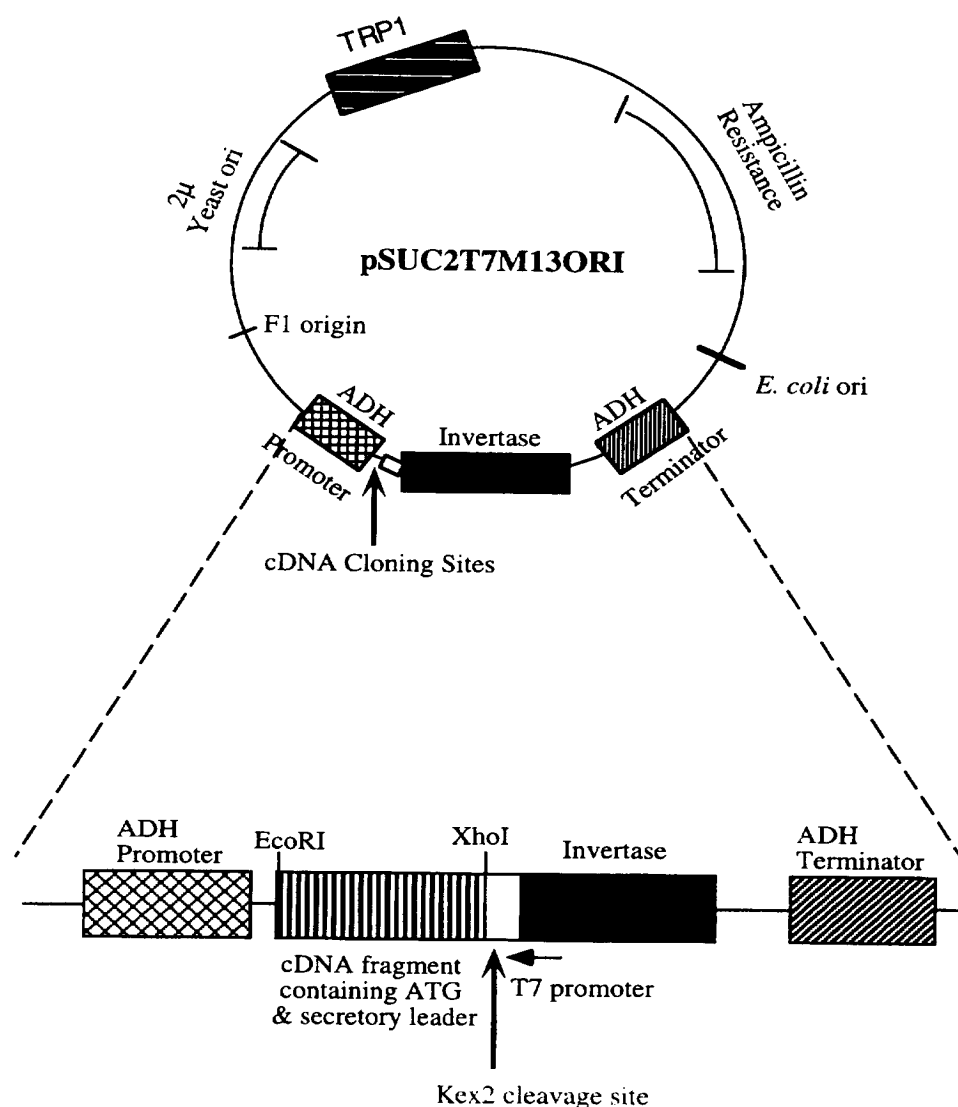


Fig. 1. The yeast signal sequence trap vector, pSUC2T7M13ORI. The invertase gene in plasmid pRB58 (Kaiser et al., 1987) (base pairs 842-2386 in GenBank accession No. V01311) was amplified using Vent polymerase (New England Biolabs, Beverly, MA, USA) and ligated to *Eco*RI and *Xho*I cut pJG4-1 (Gyuris et al., 1993). Subsequently, the first two amino acids of mature invertase, serine and methionine, were deleted and a synthetic oligo encoding a bacteriophage T7 promoter and a hybrid human serum albumin- α factor Kex2 site (Hitzeman et al., 1990) was added. The final modification was the incorporation of a bacteriophage F1 replication origin derived from pBluescript (Stratagene, La Jolla, CA, USA). This vector was used in conjunction with the yeast strain YT455 (Kaiser et al., 1987) which was converted to *TRP*⁻ (Alani et al., 1987) and then to *URA*⁻ (Boeke et al., 1987).

The *Saccharomyces cerevisiae* enzyme invertase is an essential protein in certain metabolic environments. If sucrose or raffinose is provided as the sole source of carbon, yeast must secrete invertase to grow (Carlson et al., 1983). Because of this convenient genetic selection,

invertase has been widely used to measure the effects of mutations within signal sequences (Kaiser et al., 1987; Perlman et al., 1986; Ngsee et al., 1989), to identify amino acid sequences that mediate localization to yeast vacuoles (Klionsky et al., 1988; Tague et al., 1990;

Saalbach et al., 1991) or to mitochondria (Emr et al., 1986), and to monitor the amounts of human proteins expressed for commercial purposes (Hitzeman et al., 1990). As a whole, these results demonstrated two important points: the genetic selection for invertase secretion is very sensitive, and invertase itself is quite tolerant of amino-terminal extensions. We describe here a method which uses the invertase genetic selection to isolate efficiently and in high throughput, genes encoding secreted proteins. An effective but less efficient method of trapping signal sequences was first published by Tashiro et al. (1993). Another publication using a method similar to the one we describe and with similar results has also been published (Klein et al., 1996). Here we present data regarding the selection efficiency, cloning biases, and sequences of a novel human chemokine and a novel G-protein-coupled receptor discovered using the system.

2. Results

2.1. Vector design and testing the invertase selection

The vector pSUC2T7M13ORI (Fig. 1) was designed to identify sequences in cDNA clones that mediate transport of proteins into the secretion pathway. The vector carries an invertase gene (*SUC2*) lacking both its initiating methionine codon and its signal peptide. Transcription of invertase is mediated by the yeast alcohol dehydrogenase promoter, but the invertase translation and secretion defect renders the plasmid unable to rescue the growth of a *SUC2*[−] yeast strain on YPR plates. *SUC2*[−] yeast transformed with pSUC2T7M13ORI, and placed under invertase selection produce no visible colonies appear even after 7 days of growth. When the native invertase signal sequence and initiator methionine are inserted into the cloning site, transformed *SUC2*[−] yeast grow as well on raffinose as *SUC2*⁺ yeast, with a plating efficiency of 100% after 3 days. By contrast, *SUC2*[−] yeast carrying the vector modified to encode only an initiating methionine residue upstream of invertase (i.e., capable of producing only cytoplasmic invertase) produced visible colonies only after 8 days, and with a plating efficiency of just 2%. These differences in growth rates and plating efficiencies suggested that the growth selection would be effective at isolating those clones in a cDNA library containing natural signal sequences.

2.2. Isolation of cDNA clones encoding secreted proteins

A cDNA library was constructed with mRNA isolated from activated human PBMC. This cell source was chosen as a test since it expresses a variety of known secreted cytokines, hormones, growth factors and cell

surface receptors. The library was transformed into *SUC2*[−] yeast, and plasmids were isolated from colonies which survived the invertase selection.

Two separate selections gave qualitatively similar results. In selection one, 1.9×10^6 transformed yeast (i.e., *TRP*⁺) were subjected to invertase selection, and a total of 147 *SUC2*⁺ clones were recovered. Of these, 136 (92%) corresponded to database entries of well-characterized genes and proteins. The incidence of signal sequences among this set provided a measure of the fidelity of the invertase selection, the true positive rate. Among this set, 83% of the cDNAs encoded proteins known to contain signal sequences. A representative number of these were further analyzed to verify that they were fused in-frame to invertase, consistent with their signal sequences mediating the secretion of invertase. In the second run, 4.8×10^6 *TRP*⁺ yeast were selected, 457 *SUC2*⁺ clones were analyzed, and 424 (93%) were informative. Among this set, 83% again encoded proteins known to contain signal sequences. Table 1 provides a combined listing of the known genes identified from these selections. The list includes secreted proteins, such as interleukins and chemokines, type 1 and multipass transmembrane proteins, and proteins found in intracellular organelles. Significantly, the list also includes type 2 transmembrane proteins such as Fas ligand. In common with other type 2 membrane proteins (Parks, 1996) Fas ligand has no discernable N-terminal signal sequence. Nevertheless, the molecule was isolated by the yeast signal sequence trap by virtue of its being fused to invertase via a serine residue at position 153. Fas ligand must therefore possess an uncharacterized motif within its sequence (before residue 153) which is capable of mediating secretion in yeast.

A minority (17%) of isolated clones were false positives, i.e., clones encoding proteins known not to be secreted. Most of these false positives encoded an initiating methionine adventitiously followed by an amino acid sequence that mimicked a signal sequence. These mimic sequences were found to be derived either from 3'-untranslated regions or from coding regions, with the latter often employing a reading frame different from the one used in vivo (data not shown). The presence of false positives is not surprising in light of the observation that 20% of randomly generated peptides can function as signal sequences in yeast if an initiating methionine is provided (Kaiser et al., 1987). Perhaps more surprising is that, despite this, we observe a greater than 80% specificity for signal sequences with the yeast signal sequence trap system.

2.3. Extent of recovery of cDNA clones encoding signal sequences

Given the known patterns of gene expression in PBMC, the list of genes isolated by the yeast signal

Table 1
Secreted proteins identified in a human PBMC library

Accession No.	Protein	Accession No.	Protein
Extracellular: cytokines		M97720	T cell receptor V α
M13982	IL-4	M15565	T cell receptor V α
X04688	IL-5	68696	T cell receptor V α
X17543	IL-9	07294	T cell receptor V β
X13274	γ -IFN	37797	RP105 (homolog)
X10394	Tumor necrosis factor		
X01393	Lymphotoxin		
U32659	IL-17 (homolog)		
Extracellular: chemokines		Cell surface: type II transmembrane proteins	
Z11686	IL-8	U11276	NKR-P1a protein
J04130	Act-2	U08137	Fas ligand
X72755	HuMig	X67878	CD40 ligand
X64885	RANTES	M14766	Fc epsilon RII (CD23)
X14768	MCP-1	K01144	MHC class II invariant chain (CD74)
X71087	MCP-3	L07555	Activation inducer molecule (CD69)
M23178	MIP-1 α	H11808	Activation inducer molecule (CD69) (homolog)
M24110	MIP-1 α homolog-2		
X72755	HuMig (homolog)		
Cell surface: type I transmembrane proteins		Cell surface: multi-pass transmembrane proteins	
M14362	LFA-2 (CD2)	X71635	Neuropeptide Y-like receptor
M23461	Leukocyte Common Antigen (CD45)	X07982	LIMP (CD63)
M33195	Fc RI (CD64)	U33447	Putative G-coupled protein receptor (homolog)
L15006	CTLA-4		
Z22968	M130 Antigen	Intracellular: organelle proteins	
U03397	4-1BB Receptor	U35237	Tryptase 2
D11086	IL-2 Receptor gamma	M90696	Cathepsin S (CTSS)
S93788	Ocular melanoma associated antigen	X74104	Translocon associated protein β
S82300	β 2 microglobulin	J04071	Serine esterase
U09087	Thymopoietin β	H65908	Oligopeptide transporter (homolog)
X54890	Leukocyte common antigen related peptide		
X00457	HLA-SB α -chain	Miscellaneous:	
J00194	HLA-DR α -chain	X14584	Immunoglobulin μ -chain
M60334	HLA-DR α -chain	X57809	Immunoglobulin λ -chain
K01172	HLA-DS α -chain	M12759	Ig J chain
M81140	HLA-DQ β -chain	X17042	Hematopoietic proteoglycan core protein
X58767	T cell receptor V α	J04621	Heparan sulfate proteoglycan core protein
X58746	T cell receptor V α	D14043	MGC-24 (peanut agglutinin binding protein)
X58738	T cell receptor V α	X13694	Osteopontin
M87477	T cell receptor V α	A10156	Lysozyme
		X02530	γ -IFN inducible early response gene
		X67698	Tissue specific secretory protein (HE1)
		D28593	Mannose binding protein associated serine protease (homolog)
		U01235	Pre-pro hemorrhagic toxin (homolog)
		R16862	OTS-8 (homolog)

Proteins of known cellular location isolated by the yeast signal sequence trap from a human PBMC cDNA library. Total RNA was isolated (Chomczynski and Sacchi, 1987) from human PBMC activated by culture in ConA and PMA. PolyA⁺ mRNA, isolated with the PolyATtract mRNA isolation system (Promega, Madison, WI, USA), was converted to double-stranded cDNA using the Superscript Choice System (Gibco BRL, Gaithersburg, MD, USA) and the oligo/random primer 5'-TCCCGATTGAATTCTAGACCTGCGTCGACNNNNNN-3'. *Eco*RI adaptors were ligated onto the cDNA which was then digested with *Sa*II; 300–600 base pair fragments were ligated into pSUC2T7M13ORI pre-digested with *Eco*RI and *Xho*I. Electrocompetent DH10B cells (Gibco BRL) were transformed, and 5×10^6 colonies were pooled to prepare the amplified plasmid cDNA library (Sambrook et al., 1989). Yeasts transformed using lithium acetate were selected by growth on complete minimal plates lacking tryptophan (CMD-W) (Ausubel et al., 1995) for 3 days before being replicated onto YPR plates (1% yeast extract, 2% peptone, 2% raffinose, 2 μ g/ml Antimycin A (cat no. A-8674; Sigma, St. Louis, MO, USA)) to select for plasmids encoding signal sequences. Seven days later yeast colonies were purified by streaking onto these same plates. Individual colonies were patched onto CMD-W plates, grown overnight, and then inoculated into deep-well trays to prepare plasmid DNA essentially as described (Blanchard and Nowotny, 1994). The cDNA inserts were amplified with *Taq* polymerase (Perkin-Elmer, Norwalk, CT, USA) and sequenced utilizing the chain-termination method (Sanger et al., 1977). DNA sequences were compared with sequences in GenBank and GenPept utilizing the programs BlastN and BlastX (Altschul et al., 1990) and FastA (Pearson and Lipman, 1988).

sequence trap shown in Table 1 is certainly incomplete. The two selections were not intended to be an exhaustive search for all genes encoding signal sequences expressed by PBMC; nevertheless there were some expected genes

that were not recovered (i.e., false negatives). Possible reasons for false negatives include; the complete absence of a cDNA from the library, an insufficient sample size for the yeast selections, or the existence of poorly

expressed, poorly secreted, or inactive invertase fusions for certain genes. To address some of these questions, we measured the frequencies of 11 genes expected to be present in the starting library by hybridization to 50 000 *E. coli* colonies. γ -IFN and IL-8 were very abundant both in the starting library, at frequencies of 1:300 and 1:2000, respectively, and in the clones recovered following the selections, frequencies of 1:1.6 and 1:20, respectively. Although clones encoding GM-CSF, IL-2, IL-6, and osteopontin were present in moderate abundance in the starting library, 1:10 000–1:20 000, only osteopontin was recovered by the yeast signal sequence trap in these two selections. However, we have isolated IL-2 and IL-6 in subsequent selections. Of five genes present at low frequency in the starting library less than or equal to 1:50 000, three (IL-4, IL-5, and TNF) were isolated in these two selections and the two remaining (IL-3 and IL-10) in subsequent selections. This accounts for 10 of the 11 genes. Thus, the yeast signal sequence trap does isolate rarely expressed genes, but the particular genes isolated are partly a function of sample size.

We have not yet isolated a clone of GM-CSF, indicating that the system can exhibit cloning bias (see Section 3). One possibility, that certain structural classes of protein may be impossible to isolate using the system, is unlikely given the wide structural diversity of proteins whose genes have been selected. Another is that particular cDNA fusions of specific proteins could deleteriously affect invertase either through steric hindrance or by causing folding defects. Proteins are so diverse that it is difficult to contrive a universal solution to this problem, but as a safeguard we have incorporated a Kex2 proteolytic processing site at all fusion junctions between cDNAs and invertase. Kex2 is an endogenous yeast protease that cleaves proteins with appropriately presented pairs of basic amino acids as they pass through the Golgi (Hitzeman et al., 1990; Redding et al., 1991). Cleavage of invertase fusion proteins by Kex2 during secretion should not only reduce the possibility of enzymatically inactive fusions, but should also reduce and normalize any subtle effects of particular fusions on invertase specific activity.

Because so many clones of γ -IFN were recovered in the two selections, we had an opportunity to examine many independent γ -IFN cloning events to investigate any potential bias of the system in the selection of particular γ -IFN/invertase fusion junctions. Data presented in Fig. 2 revealed the yeast signal sequence trap to be very permissive in the position of these fusion junctions, with fusion points dispersed uniformly throughout the γ -IFN sequence. The same analysis performed on a number of IL-8 clones gave similar results, suggesting that any biases introduced by the yeast selection are more likely caused by cDNA expression issues rather than by the character of the protein fragments fused to invertase.

2.4. Description of novel homologs

Novel homologs of several genes were isolated from the two selections performed on the PBMC-derived library. The full-length protein sequences for two of these novel molecules are presented in Fig. 3. Clone H174 (Fig. 3a) is homologous to the α -chemokine HuMig (Liao et al., 1995). The sequence of H174 contains both a good predicted signal sequence and the C-X-C motif characteristic of α -chemokines. It is identical to a gene of unknown function designated β -R1, which is selectively induced by β -IFN (Rani et al., 1996).

Clone H963 (Fig. 3b) is a novel human seven transmembrane receptor. The figure shows an alignment of H963 to the platelet activating receptor (Ye et al., 1991), which is its closest human homolog with defined function. The predicted full-length open reading frame for H963 contains the seven predicted hydrophobic transmembrane segments characteristic of this receptor class. Interestingly, although the original fusion to invertase formed after residue 129 of the H963 sequence, the 5'-end of the fused cDNA fragment in the original construct did not include the authentic initiating methionine codon. Instead, the yeast utilized the methionine at residue 98 to initiate translation of the fusion, and secretion was mediated by sequences in the third transmembrane segment acting as a signal peptide. There is a predicted signal sequence cleavage site immediately after this third transmembrane segment following residue 115 (SigCleave score 3.3). H963 is not unique in being isolated using a transmembrane segment; we have isolated several other receptor molecules via this route.

3. Discussion

The yeast signal sequence trap cloning approach is rapid, simple, accurate, and identifies cDNAs encoding signals that mediate the transport and secretion of proteins diverse in structure, function, and cellular location. We have isolated clones encoding secreted proteins, type 1 transmembrane proteins, type 2 transmembrane proteins, multipass transmembrane proteins, and proteins located in intracellular organelles. A small percentage of the isolates are false positives, arising from clones adventitiously encoding peptides that mimic signal sequences.

Several papers describing effective and innovative but less efficient methods of trapping signal sequences have been published. The COS-based systems (Tashiro et al., 1993; Yokoyama-Kobayashi et al., 1995) are limited to pools of 200 clones or less per transfection, while the ES-based system (Skarnes et al., 1995) is limited by the necessity of generating and screening mammalian cell lines. Each of these methods is a screen. Our system, which is similar in many respects to the one reported

FUSION JUNCTIONS OF GAMMA-INTERFERON WITH INVERTASE

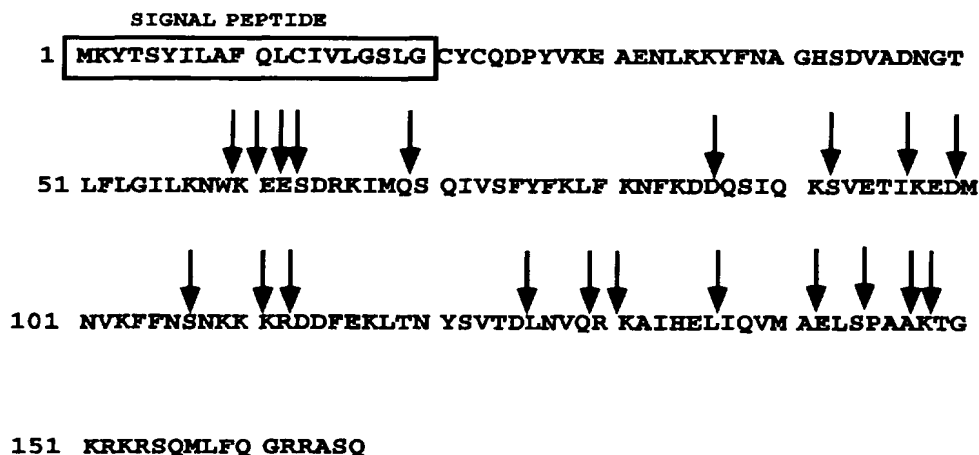


Fig. 2. Many points of fusion of γ -IFN to invertase are compatible with the genetic selection. The amino acid sequence of γ -IFN is marked with arrows to denote the positions fused to invertase. Forty-three γ -IFN clones recovered following invertase selection were sequenced, and 20 independent fusion sites were identified.

by Klein et al. (1996), is a selection, and consequently more sensitive and by design less labor intensive. We select for signal sequences among millions of transformed yeast, and positives are isolated simply by streaking colonies onto selective medium. Moreover, the yeast system identifies a more diverse repertoire of secreted sequences, as adduced by the isolation of genes encoding type 2 transmembrane proteins.

False negatives, i.e., genes encoding secreted proteins which are not identified, are a limitation of any signal-sequence cloning scheme, but may be minimal with the yeast system. The wide structural diversity of the known proteins identified using the yeast signal-sequence trap argues strongly against any broad class of proteins being totally unrepresented. Moreover, invertase activity is tolerant of many different amino terminal extensions. Kaiser et al. (1987) selected for functioning signal sequences derived from randomly sheared human genomic DNA and estimated that 20% of all open reading frames functioned well enough as signal sequences to mediate secretion of invertase, as measured by the genetic selection (Kaiser et al., 1987). In our experiments we also have observed that many diverse points of fusion between γ -IFN and invertase are readily selected (Fig. 2). Even if a particular fusion were to deleteriously effect invertase activity, it is likely that alternative fusions of the same cDNA to invertase would be fully active. The consequence may be a reduced frequency for that cDNA in the yeast selections, but it should still be recoverable.

False negatives may also be caused by a failure of the

yeast translational apparatus to initiate on heterologous mRNA sequences. However unlike *E. coli* (McCarthy and Gualerzi, 1990) and mammalian ribosomes (Kozak, 1989), yeast ribosomes do not require specific nucleotide motifs to initiate translation. The only identified criteria for translation initiation in yeast is a preference, but not an absolute requirement, for an AU-rich 5'-untranslated region (Yoon and Donahue, 1992). We suggest, therefore, that the efficiencies of heterologous translation initiation and secretion in yeast are not significant impediments to identifying clones.

The class of genes selected by the yeast signal sequence trap is enriched proteins involved in cell growth, differentiation, and intercellular communication, important cellular processes which are of high scientific and commercial interest. Since genes encoding secreted proteins represent only a fraction, perhaps 10%, of the total genome, and are encoded by rarely expressed genes, the yeast signal sequence trap is more efficient at identifying them than random sequencing. Additionally, since the yeast signal sequence trap isolates the 5'-end of mRNAs, it provides an important advantage in isolating full-length cDNAs, which can then be biologically tested.

Acknowledgement

We thank Roger Brent for many exceptionally helpful consultations, S. Herrmann and S. Clark for discussions and support, C. Kaiser for yeast strains and plasmids, R. Finley for discussions and advice on manipulating

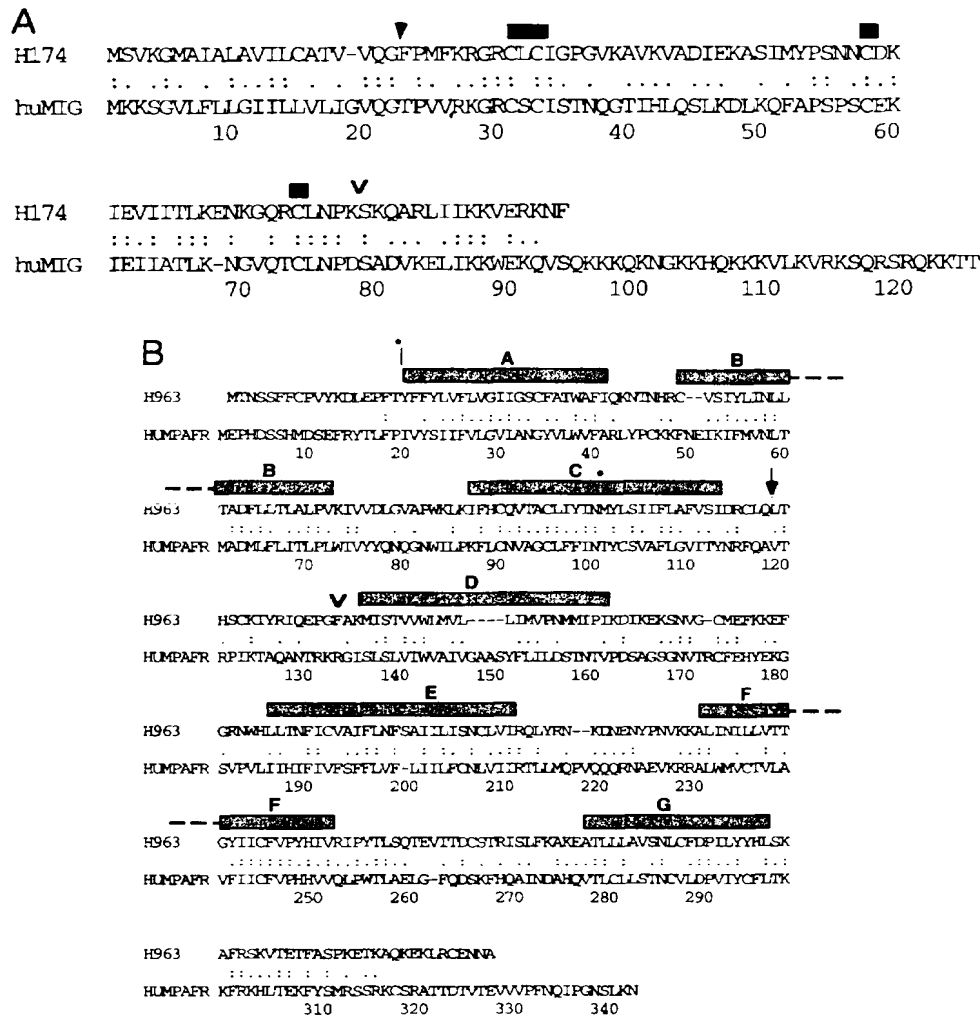


Fig. 3. Sequence alignment of novel genes identified by the Signal Sequence Trap to close homologs. Alignments were determined by FastA searches (Pearson and Lipman, 1988) against GenBank. Regions of identity are indicated by colons (:), conservative amino acid changes by single dots (.). Closed arrowheads indicate the signal sequence cleavage sites of the novel clones as determined by SigCleave (Program Manual for the Wisconsin Package, 1994); open arrowheads show the fusion junctions to invertase. (a) Comparison of the amino acid sequence of novel clone H174 to HuMig, an α -chemokine. Solid bars highlight the cysteine residues. (b) Comparison of the amino acid sequence of novel clone H963 to human platelet activating factor receptor, a seven transmembrane receptor family member. Shaded bars delineate predicted transmembrane segments. The black dot shows the methionine used for translation initiation in the yeast selection; the asterisk denotes the 5'-end of the original partial cDNA clone. The cDNA sequences are available in GenBank under accession Nos. AF002985 for H174 and AF002986 for H963.

yeast, T. Celeste for initial help with cDNA libraries, and V. Nowotny for providing his yeast DNA extraction technique prior to publication.

References

- Alani, E., Cao, L., Kleckner, N., 1987. A method for gene disruption that allows repeated use of *ura3* selection in the construction of multiply disrupted yeast strains. *Genetics* 116, 541–545.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K.J. (Eds.), 1995. *Current Protocols in Molecular Biology*. John Wiley, New York, NY, Chapter 13.
- Bernstein, H.D., Zopf, D., Freymann, D.M., Walter, P., 1993. Functional substitution of the signal recognition particle 54-kDa subunit by its *E. coli* homolog. *Proc. Natl. Acad. Sci. USA* 90, 5229–5233.
- Blanchard, M.M., Nowotny, V., 1994. High-throughput rapid yeast DNA extraction. Application to yeast artificial chromosomes as Polymerase Chain Reaction templates. *GATA* 11, 7–11.
- Blobel, G., Dobberstein, B., 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J. Cell Biol.* 67, 835–851.
- Boeke, J.D., Trueheart, J., Natsoulis, G., Fink, G.R., 1987. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol.* 154, 164–175.

- Carlson, M., Taussig, R., Kustu, S., Botstein, D., 1983. The secreted form of invertase in *Saccharomyces cerevisiae* is synthesized from mRNA encoding a signal sequence. *Mol. Cell. Biol.* 3, 439–447.
- Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* 162, 156–159.
- Emr, S.D., Vassarotti, A., Garrett, J., Geller, B.L., Takeda, M., Douglas, M.G., 1986. The amino terminus of the yeast F1-ATPase β -subunit precursor functions as a mitochondrial import signal. *J. Cell. Biol.* 102, 523–533.
- Gyuris, J., Golemis, E., Chertkov, H., Brent, R., 1993. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75, 791–803.
- Hartmann, E., Sommer, T., Prehn, S., Görlich, D., Jentsch, S., Rapoport, T.A., 1994. Evolutionary conversation of components of the protein translocation complex. *Nature* 367, 654–700.
- Hitzeman, R.A., Chen, C.Y., Dowbenko, D.J., Renz, M.E., Liu, C., Pai, R., Simpson, N.J., Kohr, W.J., Singh, A., Chisholm, V., Hamilton, R., Chang, C.N., 1990. Use of heterologous and homologous signal sequences for secretion of heterologous proteins from yeast. *Methods Enzymol.* 185, 421–440.
- Kaiser, C.A., Preuss, D., Grisafi, P., Botstein, D., 1987. Many random sequences functionally replace the secretion signal sequence of yeast invertase. *Science* 235, 312–317.
- Klein, R.D., Gu, Q., Goddard, A., Rosenthal, A., 1996. Selection for genes encoding secreted proteins and receptors. *Proc. Natl. Acad. Sci. USA* 93, 7108–7113.
- Klionsky, D.J., Banta, L.M., Emr, S.D., 1988. Intracellular sorting and processing of a yeast vacuolar hydrolase: proteinase A propeptide contains vacuolar targeting information. *Mol. Cell. Biol.* 8, 2105–2116.
- Kozak, M., 1989. The scanning model for translation: an update. *J. Cell Biol.* 108, 229–241.
- Liao, F., Rabin, R.L., Yannelli, J.R., Koniaris, L.G., Vanguri, P., Farber, J.M., 1995. Human Mig chemokine: biochemical and functional characterization. *J. Exp. Med.* 182, 1301–1314.
- McCarthy, J.E.G., Gualerzi, C., 1990. Translational control of prokaryotic gene expression. *Trends Genet.* 6, 78–85.
- Ngsee, J.K., Hansen, W., Walter, P., Smith, M., 1989. Cassette mutagenic analysis of the yeast invertase signal peptide: effects on protein translocation. *Mol. Cell. Biol.* 9, 3400–3410.
- Parks, G.D., 1996. Differential effects of changes in the length of a signal/anchor domain on membrane insertion, subunit assembly, and intracellular transport of a type II integral membrane protein. *J. Biochem.* 271, 7187–7195.
- Pearson, W.R., Lipman, D.J., 1988. Rapid and sensitive sequence comparison with FastP and FastA. *Methods Enzymol.* 183, 63–98.
- Perlman, D., Raney, P., Halvorson, H.O., 1986. Mutations affecting the signal sequence alter synthesis and secretion of yeast invertase. *Proc. Natl. Acad. Sci. USA* 83, 5033–5037.
- Program Manual for the Wisconsin Package, version 8, (August, 1994) Genetics Computer Group, 575 Science Drive, Madison, WI 53711.
- Program Manual for the EGCG Package, Peter Rice, The Sanger Centre, Hinxton Hall, Cambridge, CB10 1RQ, UK.
- Rani, M.R.S., Foster, G.R., Leung, S., Leaman, D., Stark, G.R., Ransohoff, R.M., 1996. Characterization of β -R1, a gene that is selectively induced by interferon (IFN- β) compared with IFN- α . *J. Biol. Chem.* 271, 22878–22884.
- Rapoport, T.A., Jungnickel, B., Kutay, U., 1996. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Ann. Rev. Biochem.* 65, 271–303.
- Redding, K., Holcomb, C., Fuller, R.S., 1991. Immunolocalization of Kex2 protease identifies a putative late Golgi compartment in the yeast *Saccharomyces cerevisiae*. *J. Cell Biol.* 113, 527–538.
- Saalbach, G., Jung, R., Kunze, G., Saalbach, I., Adler, K., Muntz, K., 1991. Different legumin protein domains act as vacuolar targeting signals. *Plant Cell* 3, 695–708.
- Sambrook, J.L., Fritsch, E.F., Maniatis, T. (Eds.), 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- Skarnes, W.C., Moss, J.E., Hurtley, S.M., Beddington, R.S.P., 1995. Capturing genes encoding membrane and secreted proteins important for mouse development. *Proc. Natl. Acad. Sci. USA* 92, 6592–6596.
- Tague, B.W., Dickinson, C.D., Chrispeels, M.J., 1990. A short domain of the plant vacuolar protein phytohemagglutinin targets invertase to the yeast vacuole. *Plant Cell* 2, 533–546.
- Tashiro, K., Tada, H., Heilker, R., Shirozu, M., Nakano, T., Honjo, T., 1993. Signal sequence trap: a cloning strategy for secreted proteins and type I membrane proteins. *Science* 261, 600–603.
- von Heijne, G., 1985. Signal sequences. The limits of variation. *J. Mol. Biol.* 184, 99–105.
- Walter, P., Johnson, A.E., 1994. Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* 10, 87–119.
- Ye, R.D., Prossnitz, E.R., Zou, A.H., Cochrane, C.G., 1991. Characterization of a human cDNA that encodes a functional receptor for platelet activating factor. *Biochem. Biophys. Res. Commun.* 180, 105–111.
- Yokoyama-Kobayashi, M., Sugano, S., Kato, T., Kato, S., 1995. A signal sequence detection system using secreted protease activity as an indicator. *Gene* 163, 193–196.
- Yoon, H., Donahue, T.F., 1992. Control of translation initiation in *Saccharomyces cerevisiae*. *Mol. Microbiol.* 6, 1413–1419.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>212</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution American Type Culture Collection ("ATCC")	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit 25 SEPTEMBER 1997	Accession Number 209229
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
DNA Plasmid No. PS062 In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer Sonya Barnes PCT International Division <i>SDP</i>	

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>215</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>25 SEPTEMBER 1997</u>	Accession Number <u>209300</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
DNA Plasmid No. <u>PS063</u> In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <hr/> <p>Authorized officer <u>Sonya Barnes</u> PCT International Division <u>SDB</u></p>	<p style="text-align: center;">For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <hr/> <p>Authorized officer</p>
--	--

Page 2

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>225</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection ("ATCC")	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit 09 OCTOBER 1997	Accession Number 209346
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
DNA Plasmid No. PS065 In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer Sonya Barnes PCT International Division</p>	<p style="text-align: center;">For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
---	--

Page 2

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by an applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>220</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>02 OCTOBER 1997</u>	Accession Number <u>209324</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
DNA Plasmid No. <u>PS064</u> In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <hr/> <p>Authorized officer Sonya Barnes PCT International Division</p>	<p style="text-align: center;">For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <hr/> <p>Authorized officer</p>
--	--

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

Page 2

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.



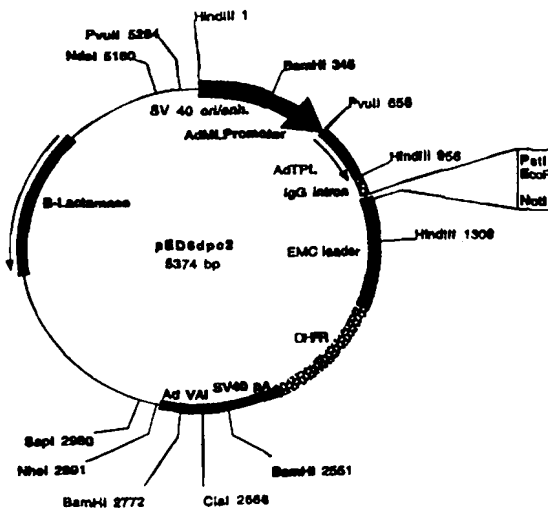
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 5/10, C07K 14/47, C12Q 1/68, A61K 38/17		A2	(11) International Publication Number: WO 98/40486
			(43) International Publication Date: 17 September 1998 (17.09.98)
(21) International Application Number: PCT/US98/04977		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 13 March 1998 (13.03.98)			
(30) Priority Data: 08/815,047 14 March 1997 (14.03.97) US 08/960,022 29 October 1997 (29.10.97) US			
(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).			
(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).		Published Without international search report and to be republished upon receipt of that report.	
(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).			

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al. (1991), NAR 19: 4485-4490.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

5

10

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

15

This application is a continuation-in-part of application Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/815,047), filed March 14, 1997.

20

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

25

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

40

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 5 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 533 to nucleotide 673;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 596 to nucleotide 673;
- 10 (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 664;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bd379_1 deposited under accession number ATCC 98361;
- 15 (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bd379_1 deposited under accession number ATCC 98361;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bd379_1 deposited under accession number ATCC 98361;
- 20 (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bd379_1 deposited under accession number ATCC 98361;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;
- 25 (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- 30 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 533 to nucleotide 673; the nucleotide sequence of SEQ ID NO:1 from nucleotide 596 to nucleotide 673; the nucleotide sequence of SEQ ID NO:1 from

nucleotide 1 to nucleotide 664; the nucleotide sequence of the full-length protein coding sequence of clone bd379_1 deposited under accession number ATCC 98361; or the nucleotide sequence of the mature protein coding sequence of clone bd379_1 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone bd379_1 deposited under accession number ATCC 98361. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:2;
 - (b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44;
 - (c) fragments of the amino acid sequence of SEQ ID NO:2; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone bd379_1 deposited under accession number ATCC 98361;
- the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 55 to nucleotide 1008;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 952 to nucleotide 1008;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 403 to nucleotide 981;

- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bp121_2 deposited under accession number ATCC 98361;
- 5 (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bp121_2 deposited under accession number ATCC 98361;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bp121_2 deposited under accession number ATCC 98361;
- 10 (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bp121_2 deposited under accession number ATCC 98361;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;
- 15 (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- 20 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 55 to nucleotide 1008; the nucleotide sequence of SEQ ID NO:3 from nucleotide 952 to nucleotide 1008; the nucleotide sequence of SEQ ID NO:3 from nucleotide 403 to nucleotide 981; the nucleotide sequence of the full-length protein coding sequence of clone bp121_2 deposited under accession number ATCC 98361; or the

25 nucleotide sequence of the mature protein coding sequence of clone bp121_2 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone bp121_2 deposited under accession number ATCC 98361. In yet other preferred

30 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 119 to amino acid 309.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
- 5 (b) the amino acid sequence of SEQ ID NO:4 from amino acid 119 to amino acid 309;
- (c) fragments of the amino acid sequence of SEQ ID NO:4; and
- (d) the amino acid sequence encoded by the cDNA insert of clone bp121_2 deposited under accession number ATCC 98361;

10 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 119 to amino acid 309.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 15 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 52 to nucleotide 639;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID
- 20 NO:5 from nucleotide 1 to nucleotide 308;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bp646_10 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the
- 25 cDNA insert of clone bp646_10 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bp646_10 deposited under accession number ATCC 98361;
- (g) a polynucleotide encoding the mature protein encoded by the
- 30 cDNA insert of clone bp646_10 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

5 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 52 to nucleotide 639; the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 308; the nucleotide sequence of the full-length protein coding
10 sequence of clone bp646_10 deposited under accession number ATCC 98361; or the nucleotide sequence of the mature protein coding sequence of clone bp646_10 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone bp646_10 deposited under accession number ATCC 98361. In yet other preferred
15 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 86.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising
20 a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 86;

25 (c) fragments of the amino acid sequence of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone bp646_10 deposited under accession number ATCC 98361;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence
30 of SEQ ID NO:6 from amino acid 1 to amino acid 86.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 134 to nucleotide 1183;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 191 to nucleotide 1183;
- 5 (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 763;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cf50_1 deposited under accession number ATCC 98361;
- 10 (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cf50_1 deposited under accession number ATCC 98361;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cf50_1 deposited under accession number ATCC 98361;
- 15 (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cf50_1 deposited under accession number ATCC 98361;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;
- 20 (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- 25 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 134 to nucleotide 1183; the nucleotide sequence of SEQ ID NO:7 from nucleotide 191 to nucleotide 1183; the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 763; the nucleotide sequence of the full-length protein coding sequence of clone cf50_1 deposited under accession number ATCC 98361; or the nucleotide sequence of the mature protein coding sequence of clone cf50_1 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of

30

clone cf50_1 deposited under accession number ATCC 98361. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 210.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ
5 ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:8;
- 10 (b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 210;
- (c) fragments of the amino acid sequence of SEQ ID NO:8; and
- (d) the amino acid sequence encoded by the cDNA insert of clone cf50_1 deposited under accession number ATCC 98361;
- 15 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 210.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 20 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 740 to nucleotide 2245;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID
25 NO:9 from nucleotide 1 to nucleotide 463;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cw1543_3 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the
30 cDNA insert of clone cw1543_3 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cw1543_3 deposited under accession number ATCC 98361;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cw1543_3 deposited under accession number ATCC 98361;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

5 (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

10 (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 740 to nucleotide 2245; the nucleotide sequence of SEQ ID NO:9
15 from nucleotide 1 to nucleotide 463; the nucleotide sequence of the full-length protein coding sequence of clone cw1543_3 deposited under accession number ATCC 98361; or the nucleotide sequence of the mature protein coding sequence of clone cw1543_3 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of
20 clone cw1543_3 deposited under accession number ATCC 98361.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group
25 consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) fragments of the amino acid sequence of SEQ ID NO:10; and

(c) the amino acid sequence encoded by the cDNA insert of clone cw1543_3 deposited under accession number ATCC 98361;

30 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 952 to nucleotide 1074;
- 5 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 524 to nucleotide 1059;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone da389_1 deposited under accession number ATCC 98361;
- 10 (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone da389_1 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone da389_1 deposited under accession number ATCC 98361;
- 15 (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone da389_1 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity;
- 20 (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- 25 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 952 to nucleotide 1074; the nucleotide sequence of SEQ ID NO:11 from nucleotide 524 to nucleotide 1059; the nucleotide sequence of the full-length protein coding sequence of clone da389_1 deposited under accession number ATCC 98361; or the nucleotide sequence of the mature protein coding sequence of clone da389_1 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone da389_1 deposited under accession number ATCC 98361. In yet other preferred

30

embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 36.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ
5 ID NO:11.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:12;
- 10 (b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 36;
- (c) fragments of the amino acid sequence of SEQ ID NO:12; and
- (d) the amino acid sequence encoded by the cDNA insert of clone da389_1 deposited under accession number ATCC 98361;
- 15 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 36.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 20 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 78 to nucleotide 1619;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID
25 NO:13 from nucleotide 604 to nucleotide 1307;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dd71_2 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the
30 cDNA insert of clone dd71_2 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone dd71_2 deposited under accession number ATCC 98361;

- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone dd71_2 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;
- 5 (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein
- 10 of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 78 to nucleotide 1619; the nucleotide sequence of SEQ ID NO:13

15 from nucleotide 604 to nucleotide 1307; the nucleotide sequence of the full-length protein coding sequence of clone dd71_2 deposited under accession number ATCC 98361; or the nucleotide sequence of the mature protein coding sequence of clone dd71_2 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of

20 clone dd71_2 deposited under accession number ATCC 98361. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14 from amino acid 200 to amino acid 410.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ

25 ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- 30 (b) the amino acid sequence of SEQ ID NO:14 from amino acid 200 to amino acid 410;
- (c) fragments of the amino acid sequence of SEQ ID NO:14; and
- (d) the amino acid sequence encoded by the cDNA insert of clone dd71_2 deposited under accession number ATCC 98361;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14 or the amino acid sequence of SEQ ID NO:14 from amino acid 200 to amino acid 410.

In one embodiment, the present invention provides a composition comprising an
5 isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1003 to nucleotide 1350;
- 10 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 714 to nucleotide 1320;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dm221_1 deposited under accession number ATCC 98361;
- 15 (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dm221_1 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone dm221_1 deposited under accession number ATCC 98361;
- 20 (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone dm221_1 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;
- (i) a polynucleotide encoding a protein comprising a fragment of the
25 amino acid sequence of SEQ ID NO:16 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- 30 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 1003 to nucleotide 1350; the nucleotide sequence of SEQ ID NO:15 from nucleotide 714 to nucleotide 1320; the nucleotide sequence of the full-length protein

coding sequence of clone dm221_1 deposited under accession number ATCC 98361; or the nucleotide sequence of the mature protein coding sequence of clone dm221_1 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone dm221_1 deposited under accession number ATCC 98361. In yet other preferred
5 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 106.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ
10 ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:16;
- 15 (b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 106;
- (c) fragments of the amino acid sequence of SEQ ID NO:16; and
- (d) the amino acid sequence encoded by the cDNA insert of clone dm221_1 deposited under accession number ATCC 98361;
- 20 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 106.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 25 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 449 to nucleotide 1006;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID
30 NO:17 from nucleotide 1 to nucleotide 331;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dx279_1 deposited under accession number ATCC 98361;

- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dx279_1 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone dx279_1 deposited under accession number ATCC 98361;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone dx279_1 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 449 to nucleotide 1006; the nucleotide sequence of SEQ ID NO:17 from nucleotide 1 to nucleotide 331; the nucleotide sequence of the full-length protein coding sequence of clone dx279_1 deposited under accession number ATCC 98361; or the nucleotide sequence of the mature protein coding sequence of clone dx279_1 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone dx279_1 deposited under accession number ATCC 98361.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
- (b) fragments of the amino acid sequence of SEQ ID NO:18; and
- (c) the amino acid sequence encoded by the cDNA insert of clone dx279_1 deposited under accession number ATCC 98361;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 5 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 74 to nucleotide 865;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID
10 NO:19 from nucleotide 538 to nucleotide 1044;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gm243_1 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the
15 cDNA insert of clone gm243_1 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone gm243_1 deposited under accession number ATCC 98361;
- (g) a polynucleotide encoding the mature protein encoded by the
20 cDNA insert of clone gm243_1 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity;
- 25 (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions
30 to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 74 to nucleotide 865; the nucleotide sequence of SEQ ID NO:19 from nucleotide 538 to nucleotide 1044; the nucleotide sequence of the full-length protein coding sequence of clone gm243_1 deposited under accession number ATCC 98361; or the

nucleotide sequence of the mature protein coding sequence of clone gm243_1 deposited under accession number ATCC 98361. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone gm243_1 deposited under accession number ATCC 98361.

5 Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- 10 (a) the amino acid sequence of SEQ ID NO:20;
(b) fragments of the amino acid sequence of SEQ ID NO:20; and
(c) the amino acid sequence encoded by the cDNA insert of clone gm243_1 deposited under accession number ATCC 98361;

the protein being substantially free from other mammalian proteins. Preferably such
15 protein comprises the amino acid sequence of SEQ ID NO:20.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions.

Processes are also provided for producing a protein, which comprise:

- 20 (a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and
(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such
25 process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical
30 condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a schematic representation of the pED6 and pNOTs vectors used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "bd379_1"

A polynucleotide of the present invention has been identified as clone "bd379_1". bd379_1 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bd379_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bd379_1 protein").

The nucleotide sequence of bd379_1 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bd379_1 protein corresponding to the foregoing

nucleotide sequence is reported in SEQ ID NO:2. Amino acids 9 to 21 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 22, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone
5 bd379_1 should be approximately 1200 bp.

The nucleotide sequence disclosed herein for bd379_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bd379_1 demonstrated at least some similarity with sequences identified as F05256 (H. sapiens partial cDNA sequence; clone c-05b06), R60369
10 (yh04b03.r1 Homo sapiens cDNA clone 42053 5'), and W39550 (zc18g02.r1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 322706 5'). Based upon sequence similarity, bd379_1 proteins and each similar protein or peptide may share at least some activity.

15 Clone "bp121_2"

A polynucleotide of the present invention has been identified as clone "bp121_2". bp121_2 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer
20 analysis of the amino acid sequence of the encoded protein. bp121_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bp121_2 protein").

The nucleotide sequence of bp121_2 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the
25 predicted amino acid sequence of the bp121_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 287 to 299 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 300, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone
30 bp121_2 should be approximately 4175 bp.

The nucleotide sequence disclosed herein for bp121_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bp121_2 demonstrated at least some similarity with sequences identified as AA261860 (zs18g12.s1 NCI CGAP_GCB1 Homo sapiens cDNA clone

IMAGE:685606 3'), AA478628 (zv19g09.s1 Soares NhHMPu S1 Homo sapiens cDNA clone 754144 3' similar to WP D1022.1 CE02575 UBIQUITIN-CONJUGATING ENZYME), H43996 (yo70h10.r1 Homo sapiens cDNA clone 183331 5'), N20622 (yx46f08.r1 Homo sapiens cDNA clone 264807 5'), N34063 (yx78a05.r1 Homo sapiens cDNA clone 267824 5' 5 similar to D82419 similar to none), N57554 (yy81e07.s1 Homo sapiens cDNA clone 279972 3'), U23517 (Caenorhabditis elegans cosmid D1022), W19342 (zb90c09.s1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 310864 3' similar to WP D1022.1 CE02575 UBIQUITIN-CONJUGATING ENZYME), and W81357 (zd86c08.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 347534 3' similar to WP D1022.1 CE02575 10 UBIQUITIN-CONJUGATING ENZYME). The predicted amino acid sequence disclosed herein for bp121_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bp121_2 protein demonstrated at least some similarity to sequences identified as U23517 (similar to ubiquitin conjugating enzyme [Caenorhabditis elegans]) and W05315 (Ubiquitin 15 conjugating enzyme 9). Based upon sequence similarity, bp121_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the bp121_2 protein sequence centered around amino acid 110 of SEQ ID NO:4.

20 Clone "bp646_10"

A polynucleotide of the present invention has been identified as clone "bp646_10". bp646_10 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer 25 analysis of the amino acid sequence of the encoded protein. bp646_10 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bp646_10 protein").

The nucleotide sequence of bp646_10 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the 30 predicted amino acid sequence of the bp646_10 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bp646_10 should be approximately 1800 bp.

The nucleotide sequence disclosed herein for bp646_10 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bp646_10 demonstrated at least some similarity with sequences identified as AA040456 (zk46f10.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 485899 5'), AA101294 (zn71f03.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 563645 5' similar to WP K07E3.6 CE04722 TRANSLOCATING ATPASE), AA179341 (zp48f01.s1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 612697 3'), N54113 (yz02e02.r1 Homo sapiens cDNA clone 281882 5'), T21123 (Human gene signature HUMGS02428), U63315 (Rattus norvegicus 25-Dx (25Dx) mRNA, complete cds), X99714 (S.scrofa mRNA for steroid membrane binding protein), and Y12711 (H.sapiens mRNA for putative progesterone binding). The predicted amino acid sequence disclosed herein for bp646_10 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bp646_10 protein demonstrated at least some similarity to sequences identified as U63315 (25-Dx [Rattus norvegicus]), X99714 (steroid membrane binding protein [Sus scrofa]), and Y12711 (putative progesterone binding protein). Based upon sequence similarity, bp646_10 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the bp646_10 protein sequence centered around amino acid 40 of SEQ ID NO:6.

20

Clone "cf50_1"

A polynucleotide of the present invention has been identified as clone "cf50_1". cf50_1 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. cf50_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "cf50_1 protein").

The nucleotide sequence of cf50_1 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the cf50_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 7 to 19 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 20, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cf50_1 should be approximately 1500 bp.

The nucleotide sequence disclosed herein for cf50_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cf50_1 demonstrated at least some similarity with sequences identified as H15004 (yl26c09.s1 Homo sapiens cDNA clone 159376 3'), H52859 (EST0013 Homo sapiens cDNA clone HTN-6-19), and R86003 (yp12c03.r1 Homo sapiens cDNA clone 187204 5'). Based upon sequence similarity, cf50_1 proteins and each similar protein or peptide may share at least some activity.

10

Clone "cw1543_3"

A polynucleotide of the present invention has been identified as clone "cw1543_3". cw1543_3 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. cw1543_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "cw1543_3 protein").

The nucleotide sequence of cw1543_3 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the cw1543_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cw1543_3 should be approximately 3300 bp.

The nucleotide sequence disclosed herein for cw1543_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cw1543_3 demonstrated at least some similarity with sequences identified as AA021431 (ze68f09.s1 Soares retina N2b4HR Homo sapiens cDNA clone 364169 3' similar to PIR:A55626 A55626 monocarboxylate transporter MCT2 - golden hamster), R68272 (yi06c07.s1 Homo sapiens cDNA clone 138444 3'), and U79304 (Human clone 23909 mRNA, partial cds). The predicted amino acid sequence disclosed herein for cw1543_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted cw1543_3 protein demonstrated at least some similarity to sequences identified as U62316 (monocarboxylate transporter 2 [Rattus

norvegicus]), U79304 (unknown [Homo sapiens]), and AF000240 (monocarboxylate transporter 3 [Gallus gallus]). Based upon sequence similarity, cw1543_3 proteins and each similar protein or peptide may share at least some activity.

5 Clone "da389_1"

A polynucleotide of the present invention has been identified as clone "da389_1". da389_1 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer
10 analysis of the amino acid sequence of the encoded protein. da389_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "da389_1 protein").

The nucleotide sequence of da389_1 as presently determined is reported in SEQ ID NO:11. What applicants presently believe to be the proper reading frame and the
15 predicted amino acid sequence of the da389_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone da389_1 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for da389_1 was searched against the
20 GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. da389_1 demonstrated at least some similarity with sequences identified as R46114 (yg49g06.s1 Homo sapiens cDNA clone 36151 3' similar to contains L1 repetitive element), R89713 (ym99h07.r1 Homo sapiens cDNA clone 167101 5'), Z63670 (H.sapiens CpG island DNA genomic MseI fragment, clone 89b11, forward read
25 cpg89b11.ft1a), and Z82170 (Human DNA sequence from PAC 326L13 containing brain-4 mRNA ESTs and polymorphic CA repeat). Based upon sequence similarity, da389_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of da389_1 indicates that it may contain a repetitive element.

30 Clone "dd71_2"

A polynucleotide of the present invention has been identified as clone "dd71_2". dd71_2 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer

analysis of the amino acid sequence of the encoded protein. dd71_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dd71_2 protein").

The nucleotide sequence of dd71_2 as presently determined is reported in SEQ ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dd71_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dd71_2 should be approximately 1700 bp.

The nucleotide sequence disclosed herein for dd71_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dd71_2 demonstrated at least some similarity with sequences identified as AA011156 (ze34h02.r1 Soares retina N2b4HR Homo sapiens cDNA clone 360915 5'), H64206 (EST0047 Homo sapiens cDNA clone HTN-6-41), U40719 (Rattus norvegicus S-adenosylmeth), and Z31048 (M.musculus expressed sequence tag MTEST167). The predicted amino acid sequence disclosed herein for dd71_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dd71_2 protein demonstrated at least some similarity to sequences identified as L09190 (trichohyalin [Homo sapiens]). Based upon sequence similarity, dd71_2 proteins and each similar protein or peptide may share at least some activity.

Clone "dm221_1"

A polynucleotide of the present invention has been identified as clone "dm221_1". dm221_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dm221_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dm221_1 protein").

The nucleotide sequence of dm221_1 as presently determined is reported in SEQ ID NO:15. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dm221_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dm221_1 should be approximately 2500 bp.

The nucleotide sequence disclosed herein for dm221_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dm221_1 demonstrated at least some similarity with sequences identified as AA117998 (mn06h05.r1 Beddington mouse embryonic region Mus musculus cDNA clone 537177 5'), AA164251 (zq46c05.s1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone 632744 3' similar to contains Alu repetitive element), AA333321 (EST37403 Embryo, 8 week I Homo sapiens cDNA 5' end), N93607 (zb69g11.s1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 308900 3'), U14568 (**ALU WARNING: Human Alu-Sb subfamily consensus sequence), U57007 (Human Ya5 subfamily Alu sequence), W20519 (zb26g03.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 303220 5'), and W25502 (zb69g11.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 308900 5'). The predicted amino acid sequence disclosed herein for dm221_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dm221_1 protein demonstrated at least some similarity to sequences identified as S58722 (X-linked retinopathy protein {C-terminal, clone XEH.8c} [human, Peptide Partial, 100 aa] [Homo sapiens]). Based upon sequence similarity, dm221_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of dm221_1 indicates that it may contain an Alu repetitive element.

Clone "dx279_1"

A polynucleotide of the present invention has been identified as clone "dx279_1". dx279_1 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dx279_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dx279_1 protein").

The nucleotide sequence of dx279_1 as presently determined is reported in SEQ ID NO:17. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dx279_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dx279_1 should be approximately 1300 bp.

The nucleotide sequence disclosed herein for dx279_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dx279_1 demonstrated at least some similarity with sequences identified as AA255685 (zs22e05.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone 685952 5'), R46317 (yj53g03.r1 Homo sapiens cDNA clone 152500 5'), and R67743 (yi28d02.r1 Homo sapiens cDNA clone 140547 5'). Based upon sequence similarity, dx279_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the dx279_1 protein sequence centered around amino acid 70 of SEQ ID NO:18.

Clone "gm243_1"

A polynucleotide of the present invention has been identified as clone "gm243_1". gm243_1 was isolated from a human adult uterus cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. gm243_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "gm243_1 protein").

The nucleotide sequence of gm243_1 as presently determined is reported in SEQ ID NO:19. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the gm243_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone gm243_1 should be approximately 3500 bp.

The nucleotide sequence disclosed herein for gm243_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. gm243_1 demonstrated at least some similarity with sequences identified as H39507 (yo54c09.r1 Homo sapiens cDNA clone 181744 5'). Based upon sequence similarity, gm243_1 proteins and each similar protein or peptide may share at least some activity.

Deposit of Clones

Clones bd379_1, bp121_2, bp646_10, cf50_1, cw1543_3, da389_1, dd71_2, dm221_1, dx279_1 and gm243_1 were deposited on March 13, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98361, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* 19: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone

bd379_1

bp121_2

Probe Sequence

SEQ ID NO:21

SEQ ID NO:22

	bp646_10	SEQ ID NO:23
	cf50_1	SEQ ID NO:24
	cw1543_3	SEQ ID NO:25
	da389_1	SEQ ID NO:26
5	dd71_2	SEQ ID NO:27
	dm221_1	SEQ ID NO:28
	dx279_1	SEQ ID NO:29
	gm243_1	SEQ ID NO:30

10 In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite) (Glen Research, cat. no. 10-1953)).

15 The design of the oligonucleotide probe should preferably follow these parameters:

- (a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;
- (b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

20 The oligonucleotide should preferably be labeled with $g^{-32}P$ ATP (specific activity 6000 Ci/mmol) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately $4e+6$ dpm/pmol.

25 The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 μ l of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 μ g/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100

µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

- 5 The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably
- 10 incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to
- 15 visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

- 20 Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated
- 25 herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also
- 30 be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form

of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

5 The present invention also provides genes corresponding to the cDNA sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which the cDNA sequences are derived and any contiguous regions of the genome necessary for the regulated expression of such genes, including but not limited to coding sequences, 5' and 3' untranslated regions, alternatively
10 spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic
15 materials.

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and
20 transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence
25 identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more
30 (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or

polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided
5 herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides .

10 The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of
15 stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) [†]	Hybridization Temperature and Buffer [†]	Wash Temperature and Buffer [†]
5	A	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
	B	<50	T _B *; 1xSSC	T _B *; 1xSSC
	C	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
	D	<50	T _D *; 1xSSC	T _D *; 1xSSC
	E	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
	F	<50	T _F *; 1xSSC	T _F *; 1xSSC
10	G	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
	H	<50	T _H *; 4xSSC	T _H *; 4xSSC
	I	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
	J	<50	T _J *; 4xSSC	T _J *; 4xSSC
	K	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
	L	<50	T _L *; 2xSSC	T _L *; 2xSSC
15	M	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
	N	<50	T _N *; 6xSSC	T _N *; 6xSSC
	O	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
	P	<50	T _P *; 6xSSC	T _P *; 6xSSC
	Q	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
	R	<50	T _R *; 4xSSC	T _R *; 4xSSC

[†]: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

[†]: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH₂PO₄, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6(log₁₀[Na⁺]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na⁺] is the concentration of sodium ions in the hybridization buffer ([Na⁺] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds.,
5 John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or
10 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an
15 expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably
20 linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the
25 protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

30 Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

5 methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

10 The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

20 The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

30 Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which

the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays

for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured
5 by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter
10 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bertagnolli et al., *J. Immunol.* 145:1706-1712, 1990; Bertagnolli et al., *Cellular Immunology* 133:327-341, 1991; Bertagnolli, et al., *J. Immunol.* 149:3778-3783, 1992; Bowman et al., *J. Immunol.* 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node
15 cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ , Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic
20 cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., *J. Exp. Med.* 173:1205-1211, 1991; Moreau et al., *Nature*
25 336:690-692, 1988; Greenberger et al., *Proc. Natl. Acad. Sci. U.S.A.* 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991;
30 Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient
5 by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells
10 express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function
15 (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides.
20 For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used
25 to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II
30 molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

10 The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, *Immunologic studies in Humans*); Herrmann et al., *Proc. Natl. Acad. Sci. USA* 78:2488-2492, 1981; Herrmann et al., *J. Immunol.* 128:1968-1974, 1982; Handa et al., *J. Immunol.* 135:1564-1572, 1985; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988; Herrmann et al., *Proc. Natl. Acad. Sci. USA* 78:2488-2492, 1981; Herrmann et al., *J. Immunol.* 128:1968-1974, 1982; Handa et al., *J. Immunol.* 135:1564-1572, 1985; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bowman et al., *J. Virology* 61:1992-1998; Takai et al., *J. Immunol.* 140:508-512, 1988; Bertagnoli et al., *Cellular Immunology* 133:327-341, 1991; Brown et al., *J. Immunol.* 153:3079-3092, 1994.

25 Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, *J. Immunol.* 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

30 Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad. Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

5 A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

 A protein of the present invention, which induces cartilage and/or bone growth
10 in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of
15 congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

 A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce
20 differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

25 Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and
30 other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. *De novo* tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce
5 differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in
10 the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve
15 tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present
20 invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of
25 non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac)
30 and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting
5 differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described
10 in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in:
Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year
15 Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related
20 activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals
25 and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin- β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example,
30 United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., *Endocrinology* 91:562-572, 1972; Ling et al., *Nature* 321:779-782, 1986; Vale et al., *Nature* 321:776-779, 1986; Mason et al., *Nature* 318:659-663, 1985; Forage et al., *Proc. Natl. Acad. Sci. USA* 83:3091-3095, 1986.

5

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells.

- 10 Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses
- 15 against the tumor or infecting agent.

- A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population
- 20 of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

- Assays for chemotactic activity (which will identify proteins that induce or prevent
- 25 chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene
- 30 Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. *J. Clin. Invest.* 95:1370-1376, 1995; Lind et al. *APMIS* 103:140-146, 1995; Muller et al. *Eur. J. Immunol.* 25: 1744-1748; Gruber et al. *J. of Immunol.* 152:5860-5867, 1994; Johnston et al. *J. of Immunol.* 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

10 The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 15 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

30 The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 5 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in 10 the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat 15 inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting 20 from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major 25 roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

30 The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

5 E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to
10 their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention
15 encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue
20 in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and
25 polynucleotides of the present invention encoding such proteins, can be used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the
30 cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

10 Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

20 Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen

5 in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

10 A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term

15 "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11,

20 IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention,

25 or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

30 A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone.

Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μ g to about 100 mg (preferably about 0.1mg to about 10 mg, more preferably about 0.1 μ g to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting
5 and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When
10 administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also
15 optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the
20 developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular
25 application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins
30 or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns.

- 5 In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, 10 ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 15 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

20 In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- α and TGF- β), and insulin-like growth factor (IGF).

25 The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering 30 various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline
5 labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without
10 limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

15 Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM

(iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1117 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTTACCTCTC TG TAGCTATG TTTTCCCATG GTTGTTTTAA GAGGAATAGA AGAAAGGAAA	60
ACAGCTTG TG AAAATCCTAA GTTGATTTG AACGAGCAAG CTGTGTTTCC TCATTA AAAC	120
ATTTATTT CG CATTGATGG TCCATAACTG CCCATTTACC TCAGGATGCC TCCATATGAT	180
GAAAATAAGA ACAGAGTTGA AAGAAGTCTC CATAAACACA ACGCACATTG GCAAATGTCA	240
TATTCTTG TT CCTTAAGGGA TTAGAGAACA CTTTCTTCTC TTTGTCTTTG CCCCCAAAGT	300
AAAAGCTATA AGCTTTTATA ATTAAATAAT AAGACTGAAT AACCATAAGC GCAAATAATA	360
TGTAGTATTA TGAGAAATAC TGGGAAAAAG GACACTTACT GTGTGACTTA AATTGATTAA	420
AGGGTTATTC AGTTCAACTC TCTTGAATCT AATTAGTATT TTTGTGTCAT TTATTATTAT	480
AGGGCACACA TTTTTFACAT TTGATTTAAC TTGACCAAAA TTAAATGAGC AAATGTTTAT	540
TGCTATGTCC ATTGTTTTCC TTTCTCTGTC ACTGTTAAAA AGAGGAGCCA TGGCTTCTGC	600
TTCTTCTGTG TATTCTCCAT TAGACCTTCT TCATCCACCC TCTTCCCCAT CCCTTTCAGC	660
TCTGAAGGGT CTATAAATGA AAGTGGGTAC CAACTGATTC AATAGGACTT ATATCTTACC	720
AAATAACGTT TTATTGTCTT TGTTCATGT ATTTGCAGAG AAATTGTAAG TATCTTTAAA	780
ACCAATTAAC AAAGCCCTGT GGGTCTTTCA ATCAAGACCT TTGTAAACAT CTCTACTAGC	840
CCATACTCCC CCAAACCTTCT TGCACATGGT AGAAGATGAC ATTAAATAAA GCACATTATA	900
AGGTGCAATG AGCTTTATTC TAAAAATATT GTCTGGATGT GAAAGTAAGT TCTTGTTTCA	960
AAAATGTTAT TAGTAAATG TTATTAGATT AAAATTATGG AGTAAGCATT TGGCAAAC TG	1020
ATTGACTCTT CACTGGAAAG ACCAGGCTTT TTAGGACACA TTTCTGTTCA TGCTTAAGGT	1080
CAGAAGTCAA TCAAAGGCAA CCAAAAAAAAA AAAAAA	1117

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Phe	Ile	Ala	Met	Ser	Ile	Val	Phe	Leu	Ser	Leu	Ser	Leu	Leu	Lys
1				5				10					15		
Arg	Gly	Ala	Met	Ala	Ser	Ala	Ser	Ser	Val	Tyr	Ser	Pro	Leu	Asp	Leu
			20				25					30			
Leu	His	Pro	Pro	Ser	Ser	Pro	Ser	Leu	Ser	Ala	Leu	Lys	Gly	Leu	
		35				40					45				

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4078 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTGGAGAGGA GGTGGCAGCG GCCCGGGAGG CCGGAGCCAA GCCAGCGACC CACCATGGAG	60
ACCCGCTACA ACCTGAAGAG TCCGGCTGTT AAACGTTTAA TGAAAGAAGC GGCAGAATTG	120
AAAGATCCAA CAGATCATTA CCATGCGCAG CCTTTAGAGG ATAACCTTTT TGAATGGCAC	180
TTACAGGTTA GAGGGCCCCC AGACTCCGAT TTTGATGGAG GAGTTTATCA CGGGCGGATA	240
GTACTGCCAC CAGAGTATCC CATGAAACCA CCAAGCATTA TTCTCCTAAC GGCTAATGGT	300
CGATTTGAAG TGGGCAAGAA AATCTGTTTG AGCATCTCAG GCCATCATCC TGAAACTTGG	360
CAGCCTTCGT GGAGTATAAG GACAGCATTA TTAGCCATCA TTGGGTTTAT GCCAACAAAA	420
GGAGAGGGAG CCATAGGTTC TCTAGATAAC ACTCCTGAGG AAAGAAGAGC ACTTGCCAAA	480
AAATCACAAG ATTTCTGTTG TGAAGGATGT GGCTCTGCCA TGAAGGATGT CCTGTTGCCT	540
TTAAAACTCG GAAGCGATTC AAGCCAAGCT GACCAAGAAG CCAAAGAAGT GGCTAGGCAA	600
ATAAGCTTTA AGGCAGAAGT CAATTCATCT GGAAACACTA TCTCTGAGTC AGACTTAAAC	660
CACTCTTTTT CACTAACTGA TTTACAAGAT GATATACCTA CAACATTCCA GGGTGCTACG	720
GCCAGTACAT CGTACGGACT CCAGAATTCC TCAGCAGCAT CCTTTCATCA ACCTACCCAA	780
CCTGTAGCTA AGAATACCTC CATGAGCCCT CGACAGCGCC GGGCCCAGCA GCAGAGTCAG	840

AGAAGGTTGT CTA~~CT~~TCACC AGATGTAATC CAGGGCCACC AGCCAAGAGA CAACCACACT 900

GATCATGGTG GGTCAGCTGT ACTGATTGTC ATCCTGACTT TGGCATTGGC AGCTCTTATA 960

TTCCGACGAA TATATCTGGC AAACGAATAC ATATTTGACT TTGAGTTATA ATATGGTTTT 1020

GTGACTTATG AGCTGTGACT CAACTGCTTC ATTAAACATT CTGCATTGGG TATAATCTAA 1080

GAATTGTTTA CAAAAAGATT ATTTTGTATT TACCCTTCAT TCCTTTTTTT GATCCTTGTA 1140

AGTTTAGTAT AAATATATCT AGACATTCAG ACTGTGTCTA GCAGTTACGT CCTGCTTAAA 1200

GGGACTAGAA GTCAAAGTTC CTTGTCTCAC TATTTGATCT GCTTTGCAGG GAAATAACTT 1260

GTTTTTCTC ATGTTTCATC TTCTTTTAT GTAAATTTGT AATACTTCC TATATGCCC 1320

TTTGAAATTT TTGGATAAAA GATGATGTTT TAAGTTCCAA TGAGTATTAC TAGTTACTCA 1380

ATACCACTTA TTGAGTACTC TGTTTCTACG TATGTAGAAT GTATAGGGAT AGAAGAGTTG 1440

AAAAGGGAAA GCAAACTCC TCAAGTAGCT TCCTTAAAAT GTCATTCATA GGAGATGTAC 1500

TGGAATTGCT CATTCGTGA CTTTATTTGT GTCCTAAACA TTCTTCAGTG AAAATAATTT 1560

TATTTCAGTC AAACATTTAT GAGGAAATGA GATCACATCT TTGTCACTGG ATGCTACTTG 1620

AAGAGGGAGT ACTTTGTAAC CACTTTGATA TGCTGTTATC ACCACCCCT GCCCTCTGCT 1680

GCCATAATCA CACAAATTTA AAAAGAAAGA AAACAGTCTT CCATAGATTT TTAAGGAAGA 1740

AAGGGCCCAA GCCAGGAGAT CGCTTGGTTT TCTTCAGAA GTTAAATGGG GGGATCTGAA 1800

GATTTGAATG TTTGGTCTGC TTTGAAATGT ATGTCTTTTG GGATGTATTA TATGCCTAGC 1860

TTTATAATCA GTATAAATTT TAATTATTCC AGGAATATGC ATAATATTGA AATATTTTAT 1920

GTCTATTTT AATAGAAAAC CTCAGGGCCC AAGTAACAGT GATAGAAGTT AGAAAAACCT 1980

TTACTTAGAA TTGTCCACCT AGTCAGAGCC CAAGAAAGAA TTTTCAGTGG AAAAAATCAAT 2040

ATATAACTTA GTGCTAGCTA GCGCCACAGA CTCTAGTAGA TAATATTATC ATCATAATGG 2100

CTGGTGAAAC CATATAATCA CAGAAAAACA TTGCCTTCAG CATGTTCACT TCGCAGCACT 2160

GAGGGCACTC TTGAGGGTGT TGTTAATGAA GATTTAATTT TTAAATACAG GTGGTTCCAA 2220

GCTTTCAAAT AGGTTATGCT CCAAAGTGT TATTTGTAAG TTAATTTTTT TACAAGTCAA 2280

ACAATGTTGG AAGTGGTATT TAGGTTCTAG ATCGGTCCAC GAAAGTTAGC CCATATGTAT 2340

ATCTTGAATA GTATAGGGGA GGGTATTCAT AAAGTCCTTA TGTGGTTTTA ACTAAGTGAA 2400

ATTATGGACA AGAGAAATAA TTGTAAAATC GTCTTAAAGG AAAATTTAAT TTTTACTCCT 2460

GTTTATGGGA CATTCGTTCT ATTAACGTGC AGACACAATT TCTGTTTCA TCTGAGAGCC 2520

AGTTTTCTT TATTTCTACA TCTAAAATAA GAACATATTG TACATTATTA TATAATACAG 2580
 AATTGTCTTA AACTTTAATA AATTCGCATT TTAAAGGTGT TTACAGATTA TTTTTTATAT 2640
 CTGTAGCTGA ATTTGTTAAA GTCTAAAAAG CTCAAGGACT TTATGAAGAT CTCATTATAT 2700
 GAGGAAAATC ATAGGTTACC ATTTTATAAC TCTATTGCCA TAAGAAAATA CACTCTAAAA 2760
 TCTTGATTTG AAACATATTA GAAACCTTGA TTCAGTGCTC AGTGGTCTCC TAGTAAGAAG 2820
 TCACCGACGG TAGCGTCATA TGAGAAGAAA GAAATCCCCA CCACCTCAAC CTCTGCTGAG 2880
 ATTGTGTGCT AGGAACAGCC TTCCCTCCGT TTCCCCTCAG TCAAACCTGA GCCAGCCTCT 2940
 GGATCGATGT GATCTTATTG CATGTTTCCA TGGGGTGTAC CTATACTTTA AGCCAATCCT 3000
 GCTGCATTCA CTGCTAAGTT AAATAAAAAG CCAAGAAGAT TTTGCACTGT GCAGATCCTT 3060
 TGCTATCTGA CTTGCATCTC TTCCCCCACC TGTCAGCTAG CCACCTGCTT GTTTGTGTTG 3120
 GGATATTTTT TAGCACCTGA AGCACCATCT GAAAGGGCCA CCATTTTCTT CTTCCTTTTC 3180
 ATCTCACATA TGCTCCCTAA AAATCCTTAA GTTGTCAATC TGATCCCCAG TGTGAGGTTA 3240
 ATGAGCAAAA TTGGTCTTTG GGGCCCTTTT TGTCCAAGCC CCACTGAAAG GCCTCTTCAG 3300
 AAAACTATTA TCTTTAAAGC CCTACTTTAA CTCCTTAATT CCAGCATACA GCTAAAACTG 3360
 GATGTATATT CTGGCAAGTA AAGGCTGAGG ACTCCTCTTT AATCCTCAGA TCTAGATAAC 3420
 TCATGACATT TTATTTGACC AACATAGCAC ATGATGAGAT ATCAAGGTAA TTAAAATAGC 3480
 ATGCTTGAAA AAAAAATACG TAATCTGTTT CACCTGTAAC TGTTTAAGCC AATAAACTTT 3540
 TCAAAATTTA TGTAATGTGG GGCTTTTATG TAGCACTTTA CGTTTTCATG CTGCTTATTG 3600
 TTTTATTCTA CTGAAAAAAA TGAATTTCAA GATTCTCAAC TTTTTTAATT TCAAAAATTG 3660
 TTTATTGTTT TGACTATAGG AATACAAAAT TTCCTATTTT GGGAGAATAA GAACTCTTTT 3720
 TGTCATTTTT GGCTATGAAT AAACCTTCTG GTCTTTTGAG ACCACCCATT TTTATAGATC 3780
 AGAATCAGAA AACAGGTAAA CCTCACTCAC ACATTTGGAC TCATTTGAAC AAAAATCTAG 3840
 GCCAAAATAC TGAAAAGCCT ATGTGTTTTT TTAATTGGAA GTATATGTAA GGTTAATGCA 3900
 TTTAGTGAAC GTGACTAACA AAGACTAATG TGCACATTAA CAGATGTACT TTTTAAGGTT 3960
 TTATGGGAGG CTGTGCATTG CTCAAAGCT GTTGGGAACG CCTTCTGAAC AGTTGCCTTC 4020
 AGAACTAGTT TGAGCTGCTC AATAAAACCA GTGACTTTAC TCATAAAAAA AAAAAAAA 4078

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 318 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Glu	Thr	Arg	Tyr	Asn	Leu	Lys	Ser	Pro	Ala	Val	Lys	Arg	Leu	Met	1	5	10	15
Lys	Glu	Ala	Ala	Glu	Leu	Lys	Asp	Pro	Thr	Asp	His	Tyr	His	Ala	Gln	20	25	30	
Pro	Leu	Glu	Asp	Asn	Leu	Phe	Glu	Trp	His	Phe	Thr	Val	Arg	Gly	Pro	35	40	45	
Pro	Asp	Ser	Asp	Phe	Asp	Gly	Gly	Val	Tyr	His	Gly	Arg	Ile	Val	Leu	50	55	60	
Pro	Pro	Glu	Tyr	Pro	Met	Lys	Pro	Pro	Ser	Ile	Ile	Leu	Leu	Thr	Ala	65	70	75	80
Asn	Gly	Arg	Phe	Glu	Val	Gly	Lys	Lys	Ile	Cys	Leu	Ser	Ile	Ser	Gly	85	90	95	
His	His	Pro	Glu	Thr	Trp	Gln	Pro	Ser	Trp	Ser	Ile	Arg	Thr	Ala	Leu	100	105	110	
Leu	Ala	Ile	Ile	Gly	Phe	Met	Pro	Thr	Lys	Gly	Glu	Gly	Ala	Ile	Gly	115	120	125	
Ser	Leu	Asp	Asn	Thr	Pro	Glu	Glu	Arg	Arg	Ala	Leu	Ala	Lys	Lys	Ser	130	135	140	
Gln	Asp	Phe	Cys	Cys	Glu	Gly	Cys	Gly	Ser	Ala	Met	Lys	Asp	Val	Leu	145	150	155	160
Leu	Pro	Leu	Lys	Ser	Gly	Ser	Asp	Ser	Ser	Gln	Ala	Asp	Gln	Glu	Ala	165	170	175	
Lys	Glu	Leu	Ala	Arg	Gln	Ile	Ser	Phe	Lys	Ala	Glu	Val	Asn	Ser	Ser	180	185	190	
Gly	Asn	Thr	Ile	Ser	Glu	Ser	Asp	Leu	Asn	His	Ser	Phe	Ser	Leu	Thr	195	200	205	
Asp	Leu	Gln	Asp	Asp	Ile	Pro	Thr	Thr	Phe	Gln	Gly	Ala	Thr	Ala	Ser	210	215	220	
Thr	Ser	Tyr	Gly	Leu	Gln	Asn	Ser	Ser	Ala	Ala	Ser	Phe	His	Gln	Pro				

225 230 235 240
 Thr Gln Pro Val Ala Lys Asn Thr Ser Met Ser Pro Arg Gln Arg Arg
 245 250 255
 Ala Gln Gln Gln Ser Gln Arg Arg Leu Ser Thr Ser Pro Asp Val Ile
 260 265 270
 Gln Gly His Gln Pro Arg Asp Asn His Thr Asp His Gly Gly Ser Ala
 275 280 285
 Val Leu Ile Val Ile Leu Thr Leu Ala Leu Ala Ala Leu Ile Phe Arg
 290 295 300
 Arg Ile Tyr Leu Ala Asn Glu Tyr Ile Phe Asp Phe Glu Leu
 305 310 315

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1868 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGCGAGTTCC GGATCCCTGC CTAGCGCGGC CCAACCTTTA CTCCAGAGAT CATGGCTGCC 60
 GAGGATGTGG TGGCGACTGG CGCCGACCCA AGCGATCTGG AGAGCGGCGG GCTGCTGCAT 120
 GAGATTTTCA CGTCGCCGCT CAACCTGCTG CTGCTTGGCC TCTGCATCTT CCTGCTCTAC 180
 AAGATCGTGC GCGGGGACCA GCCGGCGGCC AGCGGCGACA GCGACGACGA CGAGCCGCCC 240
 CCTCTGCCCC GCCTCAAGCG GCGCGACTTC ACCCCCGCCG AGCTGCGGCG CTTCGACGGC 300
 GTCCAGGACC CGCGCATACT CATGGCCATC AACGGCAAGG TGTTCGATGT GACCAAAGGC 360
 CGCAAATTCT ACSGGCCCCGA RGGGCCGTAT GGGGTCTTTG CTGGAAGAGA TGCATCCAGG 420
 GGCCTTGCCA CATTTTGCCT GGATAAGGAA GCACTGAAGG ATGAGTACGA TGACCTTTCT 480
 GACCTCACTG CTGCCCAGCA GGAGACTCTG AGTGAAGGG AGTCTCAGTT CACTTTCAAG 540
 TATCATCAGC TGGGCAAAC TCTGAAGGAG GGGGAGGAGC CCACTGTGTA CTCAGATGAG 600
 GAAGAACCAA AAGATGAGAG TGCCCGGAAA AATGATTAAA GCATTTCAGTG GAAGTATATC 660
 TATTTTGTGA TTTTGCAAAA TCATTGTGTA CAGTCCACTC TGTCTTTAAA ACATAGTGAT 720

TACAATATTT AGAAAGTTTT GAGCACTTGC TATAAGTTTT TTAATTAACA TCACTAGTGA 780
CACTAATAAA ATTAACCTCT TAGAATGCAT GATGTGTTTG TGTGTCACAA ATCCAGAAAG 840
TGAAGTCAG TGCTGTAATA CACATGTAA TACTGTTTTT CTTCTATCTG TAGTTAGTAC 900
AGGATGAATT TAAATGTGTT TTTCTGAGA GACAAGGAAG ACTTGGGTAT TTCCAAAAC 960
AGGTAAAAAT CTTAAATGTG CACCAAGAGC AAAGGATCAA CTTTGTAGTCA TGATGTTCTG 1020
TAAAGACAAC AAATCCCTTT TTTTTTCTCA ATTGACTTAA CTGCATGATT TCTGTTTTAT 1080
CTACCTCTAA AGCAAATCTG CAGTGTTCCA AAGACTTTGG TATGGATTAA GCGCTGTCCA 1140
GTAACAAAAT GAAATCTCAA AACAGAGCTC AGCTGCAAAA AAGCATATTT TCTGTGTTTC 1200
TGGACTGCAC TGTTGTCCTT GCCCTCACAT AGACACTCAG ACACCCTCAC AAACACAGTA 1260
GTCTATAGTT AGGATTAAAA TAGGATCTGA ACATTCAAAA GAAAGCTTTG GAAAAAAGA 1320
GCTGGCTGGC CTAAAAACCT AAATATATGA TGAAGATTGT AGGACTGTCT TCCAAGCCC 1380
CATGTTTCATG GTGGGGCAAT GGTTATTTGG TTATTTTACT CAATTGGTTA CTCTCATTTG 1440
AAATGAGGGA GGGACATACA GAATAGGAAC AGGTGTTTGC TCTCCTAAGA GCCTTCATGC 1500
ACACCCTGA ACCACGAGGA AACAGTACAG TCGCTAGTCA AGTGGTTTTT AAAGTAAAGT 1560
ATATTCATAA GGTAACAGTT ATTCTGTTGT TATAAACTA TACCCACTGC AAAAGTAGTA 1620
GTCAAGTGTC TAGGTCTTTG ATATTGCTCT TTTGGTTAAC ACTAAGCTTA AGTAGACTAT 1680
ACAGTTGTAT GAATTTGTAA AAGTATATGA ACACCTAGTG AGATTTCAA CTTGTAATTG 1740
TGGTTAAATA GTCATTGTAT TTTCTTGTGA ACTGTGTTTT ATGATTTTAC CTCAAATCAG 1800
AAAACAAAAT GATGTGCTTT GGTCAGTTAA TAAAAATGGT TTTACCCACT AAAAAAAAAA 1860
AAAAAAAAA 1868

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Ala	Ala	Glu	Asp	Val	Val	Ala	Thr	Gly	Ala	Asp	Pro	Ser	Asp	Leu
1				5					10					15	
Glu	Ser	Gly	Gly	Leu	Leu	His	Glu	Ile	Phe	Thr	Ser	Pro	Leu	Asn	Leu
		20						25					30		
Leu	Leu	Leu	Gly	Leu	Cys	Ile	Phe	Leu	Leu	Tyr	Lys	Ile	Val	Arg	Gly
		35					40					45			
Asp	Gln	Pro	Ala	Ala	Ser	Gly	Asp	Ser	Asp	Asp	Asp	Glu	Pro	Pro	Pro
	50					55					60				
Leu	Pro	Arg	Leu	Lys	Arg	Arg	Asp	Phe	Thr	Pro	Ala	Glu	Leu	Arg	Arg
65					70					75					80
Phe	Asp	Gly	Val	Gln	Asp	Pro	Arg	Ile	Leu	Met	Ala	Ile	Asn	Gly	Lys
				85					90					95	
Val	Phe	Asp	Val	Thr	Lys	Gly	Arg	Lys	Phe	Tyr	Xaa	Pro	Glu	Gly	Pro
			100					105					110		
Tyr	Gly	Val	Phe	Ala	Gly	Arg	Asp	Ala	Ser	Arg	Gly	Leu	Ala	Thr	Phe
		115					120					125			
Cys	Leu	Asp	Lys	Glu	Ala	Leu	Lys	Asp	Glu	Tyr	Asp	Asp	Leu	Ser	Asp
	130					135					140				
Leu	Thr	Ala	Ala	Gln	Gln	Glu	Thr	Leu	Ser	Asp	Trp	Glu	Ser	Gln	Phe
145					150					155					160
Thr	Phe	Lys	Tyr	His	His	Val	Gly	Lys	Leu	Leu	Lys	Glu	Gly	Glu	Glu
				165					170					175	
Pro	Thr	Val	Tyr	Ser	Asp	Glu	Glu	Glu	Pro	Lys	Asp	Glu	Ser	Ala	Arg
			180					185					190		
Lys	Asn	Asp													
		195													

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1428 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AACCGTTGCT GGGTGTCCCA GGGCCTGAGG CAGGACGGTA CTCCGCTGAC ACCTTCCCTT

60

TCGGCCTTGA GGTTCACG CTGGTGGCCC CAGGACGTTT CGGTTCGCATG GCAGAGTGCT 120
ACGGACGACG CCTATGAAGC CCTTAGTCCT TCTAGTTGCG CTTTGTCTAT GGCCTTCGTC 180
TGTGCCGGCT TATCCGAGCA TAACTGTGAC ACCTGATGAA GAGCAAACT TGAATCATTA 240
TATACAAGTT TTAGAGAACC TAGTACGAAG TGTTCCCTCT GGGGAGCCAG GTCGTGAGAA 300
AAAATCTAAC TCTCCAAAC ATGTTTATTC TATAGCATCA AAGGGATCAA AATTTAAGGA 360
GCTAGTTACA CATGGAGACG CTTCAACTGA GAATGATGTT TTAACCAATC CTATCAGTGA 420
AGAAACTACA ACTTTCCTTA CAGGAGGCTT CACACCGGAA ATAGGAAAGA AAAACACAC 480
GGAAAGTACC CCATTCTGGT CGATCAAACC AAACAATGTT TCCATTGTTT TGCATGCAGA 540
GGAACCTTAT ATTGAAAATG AAGAGCCAGA GCCAGAGCCG GAGCCAGCTG CAAAACAAAC 600
TGAGGCACCA AGAATGTTGC CAGTTGTTAC TGAATCATCT ACAAGTCCAT ATGTTACCTC 660
ATACAAGTCA CCTGTCACCA CTTTAGATAA GAGCACTGGC ATTGAGATCT CTACAGAATC 720
AGAAGATGTT CCTCAGCTCT CAGGTGAAAC TGCGATAGAA AAACCCGAAG AGTTTGAAAA 780
GCACCCAGAG AGTTGGAATA ATGATGACAT TTTGAAAAAA ATTTTAGATA TTAATTCACA 840
AGTGCAACAG GCACTTCTTA GTGACACCAG CAACCCAGCA TATAGAGAAG ATATTGAAGC 900
CTCTAAAGAT CACCTAAAC GAAGCCTTGY TCTAGCAGCA GCAGCAGAAC ATAAATTAAA 960
AACAATGTAT AAGTCCCAGT TATTGCCAGT AGGACGAACA AGTAATAAAA TTGATGACAT 1020
CGAAACTGTT ATTAACATGC TGTGTAATTC TAGATCTAAA CTCTATGAAT ATTTAGATAT 1080
TAAATGTGTT CCACCAGAGA TGAGAGAAAA AGCTGCTACA GTATTCAATA CATTAAAAAA 1140
TATGTGTAGA TCAAGGAGAG TCACAGCCTT ATTAAAAGTT TATTAAACAA TAATATAAAA 1200
ATTTTAAACC TACTTGATAT TCCATAACAA AGCTGATTTA AGCAAACTGC ATTTTTCAC 1260
AGGAGAAATA ATCATATTCG TAATTTCAA AGTTGTATAA AAATATTTTC TATTGTAGTT 1320
CAAATGTGCC AACATCTTTA TGTGTCATGT GTTATGAACA ATTTTCATAT GCACTAAAAA 1380
CCTAATTTAA AATAAAATTT TGTTTCAGGA AAAAAAAAAA AAAAAAA 1428

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Lys Pro Leu Val Leu Leu Val Ala Leu Leu Leu Trp Pro Ser Ser
1          5          10          15

Val Pro Ala Tyr Pro Ser Ile Thr Val Thr Pro Asp Glu Glu Gln Asn
20          25          30

Leu Asn His Tyr Ile Gln Val Leu Glu Asn Leu Val Arg Ser Val Pro
35          40          45

Ser Gly Glu Pro Gly Arg Glu Lys Lys Ser Asn Ser Pro Lys His Val
50          55          60

Tyr Ser Ile Ala Ser Lys Gly Ser Lys Phe Lys Glu Leu Val Thr His
65          70          75          80

Gly Asp Ala Ser Thr Glu Asn Asp Val Leu Thr Asn Pro Ile Ser Glu
85          90          95

Glu Thr Thr Thr Phe Pro Thr Gly Gly Phe Thr Pro Glu Ile Gly Lys
100         105         110

Lys Lys His Thr Glu Ser Thr Pro Phe Trp Ser Ile Lys Pro Asn Asn
115         120         125

Val Ser Ile Val Leu His Ala Glu Glu Pro Tyr Ile Glu Asn Glu Glu
130         135         140

Pro Glu Pro Glu Pro Glu Pro Ala Ala Lys Gln Thr Glu Ala Pro Arg
145         150         155         160

Met Leu Pro Val Val Thr Glu Ser Ser Thr Ser Pro Tyr Val Thr Ser
165         170         175

Tyr Lys Ser Pro Val Thr Thr Leu Asp Lys Ser Thr Gly Ile Glu Ile
180         185         190

Ser Thr Glu Ser Glu Asp Val Pro Gln Leu Ser Gly Glu Thr Ala Ile
195         200         205

Glu Lys Pro Glu Glu Phe Gly Lys His Pro Glu Ser Trp Asn Asn Asp
210         215         220

Asp Ile Leu Lys Lys Ile Leu Asp Ile Asn Ser Gln Val Gln Gln Ala
225         230         235         240

Leu Leu Ser Asp Thr Ser Asn Pro Ala Tyr Arg Glu Asp Ile Glu Ala
245         250         255

Ser Lys Asp His Leu Lys Arg Ser Leu Xaa Leu Ala Ala Ala Ala Glu
260         265         270

```

His Lys Leu Lys Thr Met Tyr Lys Ser Gln Leu Leu Pro Val Gly Arg
 275 280 285
 Thr Ser Asn Lys Ile Asp Asp Ile Glu Thr Val Ile Asn Met Leu Cys
 290 295 300
 Asn Ser Arg Ser Lys Leu Tyr Glu Tyr Leu Asp Ile Lys Cys Val Pro
 305 310 315 320
 Pro Glu Met Arg Glu Lys Ala Ala Thr Val Phe Asn Thr Leu Lys Asn
 325 330 335
 Met Cys Arg Ser Arg Arg Val Thr Ala Leu Leu Lys Val Tyr
 340 345 350

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3742 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GATCCCCACC ACACCACCAG CCCGGCCGCA CGGGGCACTG AGCCGGGTGC TGAGCACCCGG 60
 AGGCCCCGCC GAGGCCGGGA CTCAGGACCT GCAGAGAAAC GCCTCCTGAT TTTGTCTTAC 120
 AATGGAACCT AAAAAGTCGC CTGACGGTGG ATGGGGCTGG GTGATTGTGT TTGTCTCCTT 180
 CCTTACTCAG TTTTGTGTGT ACGGATCCCC ACTAGCTGTT GGAGTCCTGT ACATAGAATG 240
 GCTGGATGCC TTTGGTGAAG GAAAAGGAAA AACAGCCTGG GTTGGATCCC TGGCAAGTGG 300
 AGTTGGCTTG CTTGCAAGTC CTGTCTGCAG TCTCTGTGTC TCATCTTTTG GAGCAAGACC 360
 TGTCACAATC TTCAGTGGCT TCATGGTGGC TGGAGGCCTG ATGTTGAGCA GTTTTGCTCC 420
 CAATATCTAC TTTCTGTTTT TTTCCTATGG CATTGTTGTA GGTGCTCCAA ATATTGCTGT 480
 TTGGAGAAAT CTGCACCTTA ATAAGAGCAG TTATTTGTGA GAAAAAAAAA AGACAAGAAT 540
 ATATATGAGA TGGGTGTGAC GAAGATCCAG TCTCTCTACA TGAAGACCAG ACTGATTGCT 600
 CCAGTCTCAG AGATGAAAAC AATAAAGAGA ACTACCCCGA CGCAGGGGCT CTGGTAGAAG 660
 AGACGCGCCG CCCTCTTGGG AGCCGCAGCA GCAGAATGTA GAGGCGACCG TGCTGGTGGA 720
 CAGCGTATTG CGACCCAGCA TGGGCAACTT CAAGTCCAGG AAGCCCAAGT CCATCTTCAA 780

AGCGGAGAGC GGGAGGAGCC ACGGAGAAAG TCAGGAGACA GAGCATGTGG TATCCAGCCA 840
GTCAGAGTGT CAGGTGAGAG CAGGAACACC AGCTCATGAG AGTCCACAAA ACAATGCCTT 900
CAAGTGCCAA GAAACAGTGC GACTTCAACC AAGAATAGAC CAGAGGACTG CCATTTGCGC 960
AAAGGATGCT TTTGAAACTC GGCAGGACTT AAATGAGGAA GAAGCTGCTC AGGTGCATGG 1020
AGTCAAGGAC CCGGCGCCAG CATCAACCCA GAGCGTGCTT GCCGATGGGA CAGATTCTGC 1080
AGACCCCTCA CCAGTCCACA AAGATGGGCA GAATGAGGCC GACAGTGCAC CAGAAGACCT 1140
CCACTCTGTG GGGACCAGCA GGCTGCTCTA TCACATCACT GATGGTGATA ACCCACTGCT 1200
GTCGCCACGA TGCTCCATCT TCAGCCAAAG CCAGAGATTC AACTTAGACC CCGAGTCAGC 1260
CCCATCTCCA CCCAGCACTC AGCAGTTTAT GATGCCGCGG AGTTCTTCAC GCTGCAGCTG 1320
TGGAGATGGC AAGGAGCCAC AGACCATCAC CCAGCTCACC AAGCACATCC AGAGCCTCAA 1380
GCGGAAAATT CGGAAATTTG AAGAAAAATT TGAACAAGAA AAGAAATACC GGCCTTCACA 1440
TGGTGACAAG ACTTCTAATC CTGAAGTCCT GAAATGGATG AATGATTTGG CTAAAGGTGC 1500
TAAACAGCTC AAAGAACTAA AGCTAAAGCT GTCAGAAGAA CAAGGGAGTG CTCCCAAAGG 1560
TCCACCTAGA AACCTGTTGT GTGAGCAACC CACAGTCCCC AGAGAAAATG GGAAACCGGA 1620
AGCTGCGGGC CCGGAGCCAA GCTCCTCTGG AGAAGAGACT CCAGATGCTG CCTTGACATG 1680
CCTGAAGGAG AGAAGAGAGC AACTTCCTCC CCAGGAGGAT TCTAAGGTAA CTAAGCAAGA 1740
CAAGAACCCTC ATAAAGCCGC TTTATGACCG ATACAGAATT ATCAAGCAAA TCTTGTC AAC 1800
ACCTTCCCTT ATTCCAACAA TTCAGGAGGA AGAGGACTCT GATGAAGACC GTCCACAGGG 1860
AAGCCAACAA CCTTCTTTGG CAGATCCAGC ATCTCACCTT CCTGTTGGTG ACCACCTCAC 1920
CTACTCTAAT GAGACTGAGC CTGTTAGGGC CCTTTTACCA GATGAAAAGA AAGAAGTAAA 1980
ACCACCAGCT CTCTCCATGT CTAATTTACA TGAGGCTACC ATGCCTGTAC TTCTTGACCA 2040
TCTCCGAGAA ACTAGGGCTG ACAAGAAGAG ACTGCGGAAA GCCTTAAGAG AATTTGAAGA 2100
ACAGTTTTTTT AAACAAACAG GAAGAAGTCC ACAAAGGAA GATAGGATAC CAATGGCAGA 2160
TGAGTATTAT GAATATAAGC ACATAAAAGC CAAACTGAGA CTATTAGAGG TCCTCATCAG 2220
CAAGCAAGAT GTGGCCAAAA CTATTTGAGG TTCAGGAAAT GTTATGATCA CTTTCACCCA 2280
TGATATAAAG TAAAGTTTAT TTTCCCTCTGC CATCCTTGCT AAGTAGTTTT GACACAATGA 2340
AAATGGAAGC ACTTTAGTGG TAGTATTAGC TGTTTTTAAG AAGGAATAGC AAGTTTAATT 2400
ATATACAAGG AGAAGGGATT TAAACGGGGG GAAGAATACA ACAGGTAGCC ATATAATTGG 2460

GAAAAAATTC AGTGTCTCC ATGCCAAGCA GAAAACTCAT AGTCAATACA AGTATTTTTTA 2520
AAAATGTCTA ATATTTTATC AAATCTAAAT AACATAGCTA GGACACTTGT TAGGGAAAGT 2580
TTATTTAGTA TCCAAAGACT GTTTATGTTG ATGTATGGAA AAGAGCATGA TTTTAAAAAA 2640
TCAATCATAG GAGGAAAAGA AATTCGCTTT TCAAGTAGGA AGGAATACAG CTAGCAAGAA 2700
AGCAATTTAT TTGAAACTTC TAATGGATTT TTGAGTGATA AAACATTTAC TACCTTGTC 2760
TTTAAGTCTG CTAGGCTCTC AGTACCCTAA AATAAACTAG ATTGTGTTGC TATTTTTTTT 2820
CTTTCTCTAT AAAAAATAACA CATTATTTTA TCCGTTATTT GAAATTTTAC ATTTCTGGTT 2880
ACCAAAGTTC ATTCTGATAG CATGTACTTT GTGAATTATT ATCTTTGTCT ATAAGTGACA 2940
GATGTTTATA TTAAAATAAA ATATTGTATT AAAAATTTAA AATAGGTATT TTGGATAGAT 3000
ATGTGTCTGT AGTATATAAT CTAATGTGTC CATAGTATTA TTGCTAATCT TTTGGTTTAC 3060
TATAAGATGA TATAACTATT TTTTCATTGG GAATATACAT TTTTCTTAAT GTTCCAACAT 3120
CTATACTTTG TAAAGTCAAA ACATTTCCCA TGAGCTGTAG TTATTCATCC TTCTGTACAA 3180
AATGAAAAGT TTGGAAATTG TTTGCCCTGA TACCTTGAAA AAGAAGCCAG AATATTTATT 3240
TGCTTCATCA ACTTCAGTGT ATATCATTTT GTGTTATTTT ATACGAAAAC ATGTTTATTA 3300
TTTTCATTTT TGTAAAAGGA AGTAAAAGGT CAACATTTTC TCTCATGTAC CAACCTTGTT 3360
TGTATTTCTA TTTTCTGTA ATGTTTAAGT ATGGATGTTG GAAGAAATTC AACATTCTCT 3420
TATAGTTTGG ATGGGAAGAC TATTGACTAT TTCAGAAACA GACTTATTTT AGAGGCTTAT 3480
TGTTTTCTCT GTATTTACCT AATATTTTAT AACTTTTATG AATCAGAATA ATGTCCTTCA 3540
TAAATTTGTT TAATTGAAGT CATCTACTTY TAACAGGACA GATACACAAC TATTTGAGGT 3600
TTACAAATTA CATCTTTGAT AAGGGAAATG GTTTCGTGAC ATGTACACAG TTGCTATTAA 3660
AATGTAACTC TATATATTCT ATATGATTGT AAATATTTTA TACAACAATA CAAATAAAAT 3720
ATTTTTCTAT TAAAAAATAA AA 3742

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 502 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Gly	Asn	Phe	Lys	Ser	Arg	Lys	Pro	Lys	Ser	Ile	Phe	Lys	Ala	Glu	1	5	10	15
Ser	Gly	Arg	Ser	His	Gly	Glu	Ser	Gln	Glu	Thr	Glu	His	Val	Val	Ser	20	25	30	
Ser	Gln	Ser	Glu	Cys	Gln	Val	Arg	Ala	Gly	Thr	Pro	Ala	His	Glu	Ser	35	40	45	
Pro	Gln	Asn	Asn	Ala	Phe	Lys	Cys	Gln	Glu	Thr	Val	Arg	Leu	Gln	Pro	50	55	60	
Arg	Ile	Asp	Gln	Arg	Thr	Ala	Ile	Ser	Pro	Lys	Asp	Ala	Phe	Glu	Thr	65	70	75	80
Arg	Gln	Asp	Leu	Asn	Glu	Glu	Glu	Ala	Ala	Gln	Val	His	Gly	Val	Lys	85	90	95	
Asp	Pro	Ala	Pro	Ala	Ser	Thr	Gln	Ser	Val	Leu	Ala	Asp	Gly	Thr	Asp	100	105	110	
Ser	Ala	Asp	Pro	Ser	Pro	Val	His	Lys	Asp	Gly	Gln	Asn	Glu	Ala	Asp	115	120	125	
Ser	Ala	Pro	Glu	Asp	Leu	His	Ser	Val	Gly	Thr	Ser	Arg	Leu	Leu	Tyr	130	135	140	
His	Ile	Thr	Asp	Gly	Asp	Asn	Pro	Leu	Leu	Ser	Pro	Arg	Cys	Ser	Ile	145	150	155	160
Phe	Ser	Gln	Ser	Gln	Arg	Phe	Asn	Leu	Asp	Pro	Glu	Ser	Ala	Pro	Ser	165	170	175	
Pro	Pro	Ser	Thr	Gln	Gln	Phe	Met	Met	Pro	Arg	Ser	Ser	Ser	Arg	Cys	180	185	190	
Ser	Cys	Gly	Asp	Gly	Lys	Glu	Pro	Gln	Thr	Ile	Thr	Gln	Leu	Thr	Lys	195	200	205	
His	Ile	Gln	Ser	Leu	Lys	Arg	Lys	Ile	Arg	Lys	Phe	Glu	Glu	Lys	Phe	210	215	220	
Glu	Gln	Glu	Lys	Lys	Tyr	Arg	Pro	Ser	His	Gly	Asp	Lys	Thr	Ser	Asn	225	230	235	240
Pro	Glu	Val	Leu	Lys	Trp	Met	Asn	Asp	Leu	Ala	Lys	Gly	Arg	Lys	Gln	245	250	255	
Leu	Lys	Glu	Leu	Lys	Leu	Lys	Leu	Ser	Glu	Glu	Gln	Gly	Ser	Ala	Pro	260	265	270	
Lys	Gly	Pro	Pro	Arg	Asn	Leu	Leu	Cys	Glu	Gln	Pro	Thr	Val	Pro	Arg				

[illegible]

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2061 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AATCGTCGGG AAGTGTTTTT GAGAAGTCTC GGTCGGTAAG GGAAGTCTTC CAAGTCCGTG	60
CAGCACTAAC GTATTGGCAC CTGCCTCCTC TTCGGCCACC CCCCAGATGA GGCAGCTGTG	120
ACTGTGTCAA GGGAAGCCAC GACTCTGACC ATAGTCTTCT CTCAGCTTCC ACTGCCGTCT	180
CCACAGGTGG GCTTCACTTT CGTGGAATCC TTGGGCTGCC GAGTTACACC TTAGGAATCC	240
TCTAATTTTC TTTCCACCTT TTGCACGCAC GCCAGGAGAT TTCTTTTCTT CATCTGTCCA	300
GTGAGGTAC CGTTTTTACT TCACAGGATT GTTGTGAAGA CCGAATTGCC AAGTGCAGTT	360
CCTGGCGCGG AGTAGGCAGG TCTTATAAAT ATTGGTTCAG TCTGAAGTTT ATCCTGGTTG	420
TTTCCCTTCT GATAATTTTT TAAGCACTTT TTATTGCTG GGTGTTTTCA CATACTTGAT	480
GGCCATCTGA CAGATGAGCA AGGAGGCTCA GAAGCTCAGC TTAAGATTTA AAAAAAGCA	540
GGGGGGCTAG AATTTAAATC AAGGTCTATC TGATGTCTAA GCTACCTATT CTGTTATACT	600
GCATAATACC CTTTTTATAT TATTTTTTAT ATTTAATCAG TAACATATGT AGATAGTACA	660
AAATTCAACA GATATCAAAG TGTGTTAAGT TTACCTTTCC ACCCACTTTC TCATTTTTGT	720
CTCCCCCAGT TCCTTTTGCA TTATTCCACG TATATTCTGT GCATATATAC ATTCATATAC	780
ATTTATCTGT ATGTGTCAGC TTCTTTTTAC ACAAATGATA CATAAACACT GTTCTGGACC	840
TTCCAACCTA GAATTACTGC AAACAGTGTC GTGATGAATT ACCTAATTCT GTGTATGTGT	900
GTATATTGGT AGAAAAAATT CCCGGAAGTA GAATTGCTAG AACAAAGATT TATGCATTTT	960
AAATATTCCT TTATTATAAA ACTAATGAAA GTAAACATGT TGGCTATGAC CACGTATGCT	1020
CTATGCTCAG TTTTCTAGA GTTGTGTATG CTTAATATAG GAGTAAGATT CTTTTAAAT	1080
GGTATATTCA TIGCCTTATT TGATTTTCAT AGTCAATCGT TTTAATTTTT CAGTCTACAT	1140
ATATAGGTGT TTGGAAGGA TATAAATATC TTCTGCTGCA TGTACCTACA GTGATAAACT	1200
CTCTCCTCCT ACATACCTTT GAGATTTTTT TTTTTTTTTT GAGACAGAGT CTCTCTCTGT	1260
CACTCAGGCT GGAGTGCAGT GGCACAGTCT GGGCTCACTG CATCCTCTGC CTACCGGGTT	1320
CAAGCAGTTC TCCTGCCTCA GCCTCTCGAG TAGCTGGGAT TACAGGCACC TGTCACCACG	1380
CCTGGCTAAT TTTTGTATTT TTAGTTGAGA CGGGGTTTCA CCATGTTGGG CAGGCTAGTC	1440
TCGAACTCCT GACCTCAAGT GATCCGCCTG CCTTGGCCTC CCACAGTGTT GGGATTACAG	1500
GTGTGAGCCA CCGTGCCTGG CCTACCTTTG AGATTTGTGA TGAGGAAACA AGAGATGAAT	1560

```

TGTATGAGAG CACTTCAAAA GATTCATGGA AAATACTTAT TTCAAAAAGA GTAGTTAATA      1620
TTACCTTATT TTTCTTATCT GCTAACCCTT TTCTTTCAAA TGCACTTAGG ACTTGCTGCT      1680
AAAACTCACT GCAAGTAAGA TACCACAAGG AGGCAGCATA GAACTGATTT TCTATACATG      1740
CTCAGGACAG TAGTTTCACT CATAGATGAA AAGTTAGAAT TTGGATTTAT TTGAAATATA      1800
TACAAATATT CAAGTATATA CATATATTCA AATAAATACA TATATGTATA TATGTGTGTA      1860
TATACACACA TACATACACA TGAATCATCA TTGCCTTCTT GAGATCTCAC CACTTTAGTC      1920
CTACTAAGAT GGGTGGTTGT TGGTTTTTTT TTGTTGTTGT TGTGTTTTTT TAAATTCCAA      1980
TCTGTATGGA ATGATACTTT AATAAAATTA TGTGCTCGGA TGTGAATAA ATGTCAAATT      2040
GCCATAAAAA AAAAAAAAAA A                                             2061

```

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 41 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

Met His Phe Lys Tyr Ser Phe Ile Ile Lys Leu Met Lys Val Asn Met
1           5           10           15

Leu Ala Met Thr Thr Tyr Ala Leu Cys Ser Val Phe Leu Glu Leu Cys
          20           25           30

Met Leu Asn Ile Gly Val Arg Phe Phe
          35           40

```

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1772 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGAAGCGGG TCCCGCAGGT CGCCACGGTT GGGGGAAACG CGGCGGACGC CGCCCCCGTC	60
CCGAAGGGGA CTCGAAAATG TACAGCCAGC GGTTCGGCAC CGTACAGCGG GAGGTTAAGG	120
GGCCACCCCC CAAAGTGGTG ATCGTGAGAT CCAAGCCTCC TAAAGGCCAA GGAGCTGAGC	180
ACCATCTAGA AAGAAATCCGA CGCAGCCATC AGAAGCATAA TGCTATTTTG GCTTCCATTA	240
AGTCAAGTGA GCGGGATCGC TTGAAAGCTG AGTGGGACCA GCACAATGAC TGCAAGATTT	300
TGGACAGCCT TGTGCGAGCA AGAATCAAGG ATGCTGTGCA AGGGTTTATC ATTAACATTG	360
AAGAAAGACG AAATAAGCTA CGTGAGCTTT TAGCATTAGA AGAAAATGAG TATTTTACAG	420
AAATGCAATT GAAGAAAGAA ACCATTGAGG AGAAAAAAGA TAGGATGAGA GAGAAAATA	480
AATTACTAAA AGAGAAGAAT GAAAAAGAGA GGCAGGATTT TGTGGCTGAA AAGCTAGACC	540
AGCAATTCAG GGAACGCTGT GAGGAGCTCC GTGTGTAATG GTTATCTATC CATCAGAAGA	600
AGGTGTGTGA GGAGCGGAAA GCACAGATTG CATTTAATGA GGAGCTGAGC AGGCAAAAGC	660
TGGTGGAAGA GCAGATGTTT TCCAACTCT GGGAGGAAGA CCGATTAGCC AAGGAAAAGC	720
GAGAAGCCCC AGAGGCGAGG AGACAGAAAG AGCTGATGGA GAACACACGC CTGGGGCTGA	780
ATGCCCAGAT CACCAGCATC AAGGCACAAA GGCAGGCGAC ACAGCTGCTG AAGGAAGAGG	840
AGGCACGCCT TGTGGAAAGT AACAACGCAC AGATTAAACA TGAGAATGAA CAGGATATGC	900
TAAAGAAACA GAAGGCAAAG CAGGAACTA GGACCATTTT GCAAAAAGCC CTACAAGAGA	960
GGATAGAACA TATTCAGCAG GAATACAGAG ACGAACAGGA CTTGAACATG AAGCTCGTGC	1020
AAAGGGCCCT TCAAGACTTA CAGGAAGAGG CAGATAAAAA GAAACAAAAA AGAGAAGATA	1080
TGATAAGAGA ACAGAAGATA TACCATAAAT ATTTGGCACA GAGACGTGAG GAAGAAAAAG	1140
CTCAGGAGAA AGAATTTGAC AGAATATTAG AGGAAGACAA GGCAAAGAAG TTGGCTGAGA	1200
AGGACAAGGA GCTGAGACTT GAAAAGGAGG CAAGGAGACA GCTTGTGGAT GAGGTCATGT	1260
GTACAAGAAA ACTTCAAGTT CAAGAAAAGT TGCAACGAGA AGCTAAAGAA CAGGAAGAAC	1320
GTGCTATGGA ACAGAAACAC ATAAATGAAA GTCTTAAAGA ACTTAACTGT GAAGAGAAGG	1380
AGAATTTTGC AAGACGCCAA CGTTTAGCCC AGGAGTACAG GAAGCAACTT CAGATGCAAA	1440
TCGCCTACCA GCAGCAGTCC CAAGAAGCAG AGAAGGAAGA GAAACGCCGA GAGTTTGAAG	1500
CAGGTGTAGC AGCAAACAAG ATGTGTTTGG ACAAGGTCCA GGAGGTCCTG TCCACCCATC	1560
AAGTGCTGCC TCAAAACATT CATCCCATGC GCAAGGCATG CCCCAGTAAG CTTCCACCGT	1620

AGTTCCGTGA GCATCAATAT ATCTTTTCTT GGTCTTTTAA TATTTTAAAC TACAGTATGC 1680
 TTGTATGCTT CTTTAACTC CTGGATAAAC TTTTCTTTTT TCCCTGAAAA AAAAAAAAAA 1740
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA 1772

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 514 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Tyr	Ser	Gln	Arg	Phe	Gly	Thr	Val	Gln	Arg	Glu	Val	Lys	Gly	Pro	1	5	10	15
Thr	Pro	Lys	Val	Val	Ile	Val	Arg	Ser	Lys	Pro	Pro	Lys	Gly	Gln	Gly	20	25	30	
Ala	Glu	His	His	Leu	Glu	Arg	Ile	Arg	Arg	Ser	His	Gln	Lys	His	Asn	35	40	45	
Ala	Ile	Leu	Ala	Ser	Ile	Lys	Ser	Ser	Glu	Arg	Asp	Arg	Leu	Lys	Ala	50	55	60	
Glu	Trp	Asp	Gln	His	Asn	Asp	Cys	Lys	Ile	Leu	Asp	Ser	Leu	Val	Arg	65	70	75	80
Ala	Arg	Ile	Lys	Asp	Ala	Val	Gln	Gly	Phe	Ile	Ile	Asn	Ile	Glu	Glu	85	90	95	
Arg	Arg	Asn	Lys	Leu	Arg	Glu	Leu	Leu	Ala	Leu	Glu	Glu	Asn	Glu	Tyr	100	105	110	
Phe	Thr	Glu	Met	Gln	Leu	Lys	Lys	Glu	Thr	Ile	Glu	Glu	Lys	Lys	Asp	115	120	125	
Arg	Met	Arg	Glu	Lys	Thr	Lys	Leu	Leu	Lys	Glu	Lys	Asn	Glu	Lys	Glu	130	135	140	
Arg	Gln	Asp	Phe	Val	Ala	Glu	Lys	Leu	Asp	Gln	Gln	Phe	Arg	Glu	Arg	145	150	155	160
Cys	Glu	Glu	Leu	Arg	Val	Glu	Leu	Leu	Ser	Ile	His	Gln	Lys	Lys	Val	165	170	175	
Cys	Glu	Glu	Arg	Lys	Ala	Gln	Ile	Ala	Phe	Asn	Glu	Glu	Leu	Ser	Arg				

180

185

190

Gln Lys Leu Val Glu Glu Gln Met Phe Ser Lys Leu Trp Glu Glu Asp
 195 200 205

Arg Leu Ala Lys Glu Lys Arg Glu Ala Gln Glu Ala Arg Arg Gln Lys
 210 215 220

Glu Leu Met Glu Asn Thr Arg Leu Gly Leu Asn Ala Gln Ile Thr Ser
 225 230 235 240

Ile Lys Ala Gln Arg Gln Ala Thr Gln Leu Leu Lys Glu Glu Glu Ala
 245 250 255

Arg Leu Val Glu Ser Asn Asn Ala Gln Ile Lys His Glu Asn Glu Gln
 260 265 270

Asp Met Leu Lys Lys Gln Lys Ala Lys Gln Glu Thr Arg Thr Ile Leu
 275 280 285

Gln Lys Ala Leu Gln Glu Arg Ile Glu His Ile Gln Gln Glu Tyr Arg
 290 295 300

Asp Glu Gln Asp Leu Asn Met Lys Leu Val Gln Arg Ala Leu Gln Asp
 305 310 315 320

Leu Gln Glu Glu Ala Asp Lys Lys Lys Gln Lys Arg Glu Asp Met Ile
 325 330 335

Arg Glu Gln Lys Ile Tyr His Lys Tyr Leu Ala Gln Arg Arg Glu Glu
 340 345 350

Glu Lys Ala Gln Glu Lys Glu Phe Asp Arg Ile Leu Glu Glu Asp Lys
 355 360 365

Ala Lys Lys Leu Ala Glu Lys Asp Lys Glu Leu Arg Leu Glu Lys Glu
 370 375 380

Ala Arg Arg Gln Leu Val Asp Glu Val Met Cys Thr Arg Lys Leu Gln
 385 390 395 400

Val Gln Glu Lys Leu Gln Arg Glu Ala Lys Glu Gln Glu Glu Arg Ala
 405 410 415

Met Glu Gln Lys His Ile Asn Glu Ser Leu Lys Glu Leu Asn Cys Glu
 420 425 430

Glu Lys Glu Asn Phe Ala Arg Arg Gln Arg Leu Ala Gln Glu Tyr Arg
 435 440 445

Lys Gln Leu Gln Met Gln Ile Ala Tyr Gln Gln Gln Ser Gln Glu Ala
 450 455 460

Glu Lys Glu Glu Lys Arg Arg Glu Phe Glu Ala Gly Val Ala Ala Asn
 465 470 475 480

Lys Met Cys Leu Asp Lys Val Gln Glu Val Leu Ser Thr His Gln Val
485 490 495

Leu Pro Gln Asn Ile His Pro Met Arg Lys Ala Cys Pro Ser Lys Leu
500 505 510

Pro Pro

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2555 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AACGGGAAAT GCTCTGTATC CAAATGCCAG CTGAGATCAT TAGCCACGCT GAAATTCGTT	60
GACTCGTACT GTGCTAAGGT TGCTGCTAGA CGTTGTTAAC AGAGTTAGTG GTCTTTGGTT	120
ATGGTGTTTA GCAGCCCCAC TTATCTGTTC CATACTACCA GTGAATAGTT TAAATTCAT	180
GTTGCACCTA TAATTTATCC CACATAAATA ATTCAGGCTA TTTATTTTGG CATTCAATTG	240
ATTTTTTTCT CACTTTAAAA AACTGAGGTA TGGAGACTGG TGCTAGCAAC ACGGGATTGG	300
CTAACGCATC CTCTTGCTGT TCCCGGTGTT TGGGCCTTGC CTGTGACAGT GGGAAAAAAA	360
ATGGCCTTGC TGTGCTACAA CCGGAGCTGC GGTCAGCGCT TCGATCCTGA GACCAATTCC	420
GACGATGCTT GCACATATCA CCCAGGCGTT CCAGTCTTTC ACGATGCATT AAAGGGTTGG	480
TCTTGCTGTA AGAGAAGAAC AACTGATTTT TCTGATTTCT TAAGCATTGT AGGCTGTACA	540
AAAGGTAGAC ATAATAGTGA GAAGCCACCT GAGCCAGTCA AACCTGAGGT CAAGACTACT	600
GAGAAGAAGG AACTATCTGA ATTAAAACCA AAATTTCAAG AACACATTCA AGCCCCTAAG	660
ACAGTAGACG CGATAAAAAG ACCAAGCCCA GATGAACCAA TGACAAATTT GGAATTAAAA	720
ATATCTGCTT CCCTTAAAC AAGCACTTGA TAACTTAAA CTGTCATCAG GGAATGAAGA	780
AAATAAGAAA AGAAGACAAT GATGAAATTA AGATTGGGAC CTCATGTAAG AATGGAGGGT	840
GTTCAAAGAC ATATCGGGGT CTAGAGAGTC TAGAAGAAGT CTGTGTATAT CATTCTGGAG	900
TACCTATTTT CCATGAGGGG ACGAAATACT GGAGCTGTTG TAGAAGAAAA ACTTCTGATT	960

TTAATACATT CTTAGACCAA GAGGGCTGTA CAAAAGAGAA ACATGTGGAC TAAAAAAGAT 1020
GCTGGGAAAA AAGTTGTTCC ATGTAGACAT GACTGGCATC AGACTGGAGG TGAAGTTACC 1080
ATTTTCAGTAT ATGCTAAAAA CTCACCTCCA GAACTTAGCC GAGTAGAAGC AAATAGCACA 1140
TTGTTAAATG TGCATATTGT ATTTGAAGGA GAGAAGGAAT TTGATCAAAA TGTGAAATTA 1200
TGGGGTGTGA TTGATGTAAA GCGAAGTTAT GTAACATGA CTGCAACAAA GATTGAAATC 1260
ACTATGAGAA AAGCTGAACC GATGCAGTGG GCAAGCCTTG AACTGCCTGC AGCTAAAAAG 1320
CAGGAAAAAC AAAAAGACGA CACAGCAGAT TGAGTGGGAG ATGGGAGGAA GGCTATTACG 1380
TATTTTCAGAA TTTTAAATAC TGTGTGAAGT GGTGGCTTGC TGCTGTCATC TTTTGTTTTG 1440
TTGTTGTGTT ACTGAATGTG GCATTTTCAGG GTTAACATTA GGTTCCTAAA AGCCAAAGTC 1500
AGTTTATCTT TTTGTGCCTC TCATCTTTCT TTCGTGTTAT GTAAGATTGA TTATTCGTTT 1560
CTCCCTACTG GTAGGAACCA TAGTTGTGTC CTGTACTTGA AGAGGCTGAA AATAGCCCA 1620
TAACCATAAT TGCAGTATTT CTTTGTATTT CTCTGTAAAG CAAAGAAATA TTAAGGAACT 1680
TTTTTTATGT CTTTGTATTA TTCCATAATT AGTAAAGCTA ATTGTGAATG TCCAATTTTA 1740
ATGAAATGTC CAATTTTAAT CAGTTTTTTT CATGGATTG TGTTCCTACG GTACTTGAAA 1800
ATATTTAAGG AAGAGATGAA GCTCTGCAGT TTTTCTATG TGGGATGATT GCTTTTTTAA 1860
GGAGGATTAA TTCTGAGGTA GTATAGTAAG TAAAGGGGAA TATATGAATT GTTAACAAAT 1920
TAGGATTTGT TTACAACCTAC TTGAATTTTT AAATTATGTC AAAACTTACA TTACTTGCCA 1980
AGCAGTATGA TGTAAGAGTA TAGGAAACAT AAATAAGAAT ACAGAGGTAT CAATTTGATT 2040
AAAATTCACC ATTTTATAAG ACTAAGCAAT AATCTTAAAA ACCTCTTTCC TGAATATTTA 2100
AATGTGTTTG TATGGTGTTA TGACTAATTG TTACTGATTT ACAGACTAAG CCCTCTTAAA 2160
ACCTTTAGTT AAATATAAAA AGAAATTATA TATATYTKGC CTCCCTGATG GAAACTATG 2220
TAAAATTGTA GACTTAAAAG GTTTGTGGAG GCCGGGCGCG GTGGYTCATG CCTGTAATCC 2280
CAGCACTTTG GGAGGCCGAG GCGGGCAGAT CACGAGGTCA GGAGATSGAG ACCATCCTGG 2340
CTAACACGGT GAAACCCCGT CTYTAGTAGA AATACAAAAA TTAGCTGGGC GTAGTGGTGG 2400
GTGCCTGTAG TCCCAGCTAC TCAGGAGGCT GAGGCAGGAG AATGGCGTGA ACCTAAGAGG 2460
CGGAGCTGGC AGTGAGCTGA GATTGCGCCA CTGCACTCCA GCCTGGGCGA CAGAGCCAGA 2520
CTCCGACTCA AAAAAAAAAA AAAAAAAAAA AAAAA 2555

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 116 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```

Met Trp Thr Lys Lys Asp Ala Gly Lys Lys Val Val Pro Cys Arg His
 1              5              10              15

Asp Trp His Gln Thr Gly Gly Glu Val Thr Ile Ser Val Tyr Ala Lys
 20              25              30

Asn Ser Leu Pro Glu Leu Ser Arg Val Glu Ala Asn Ser Thr Leu Leu
 35              40              45

Asn Val His Ile Val Phe Glu Gly Glu Lys Glu Phe Asp Gln Asn Val
 50              55              60

Lys Leu Trp Gly Val Ile Asp Val Lys Arg Ser Tyr Val Thr Met Thr
 65              70              75              80

Ala Thr Lys Ile Glu Ile Thr Met Arg Lys Ala Glu Pro Met Gln Trp
 85              90              95

Ala Ser Leu Glu Leu Pro Ala Ala Lys Lys Gln Glu Lys Gln Lys Asp
 100             105             110

Asp Thr Ala Asp
 115
  
```

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1307 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

AGTGTGGACG GCCCACTGGG TTGGTGGTGG TGGGTGCCCC TCACAGGGCT GGAGGTGTGG      60
CCGGCCCCACT GGGTTGTGCT TTCTGCCGTA CGTCCCTTCC CATGAGGATG AGATGACCCA      120
  
```

TCTGTTGCAT CCCGGCTGCT GATAAAACAA GACCCTCGGA GCCAAGAAAC AACACTGAGT	180
TCCAGATTTC GGAAGGTTCA CGAGTGTTGC CGACACGCCC TCCCAACTGC AGACATCCTC	240
CCTGGAGGAC CTGCTGTGCT CACATGCCCC CCTGTCCAGC GAGGACGACA CCTCCCCGGG	300
CTGTGCAGCC CCCTCCCAGG CACCCTTCAA GGCCTTCCTC AGTCCCCCAG AGCCACATAG	360
CCACCGAGGC ACCGACAGGA AGCTGTCCCC GCTCCTGAGC CCCTTGCAAG ACTCACTGGT	420
GGACAAGACC CTGCTGGAGC CCAGGGAGAT GGTCCGGCCT AAGAAGGTGT GTTTCTCGGA	480
GAGCAGCCTG CCCACCGGGG ACAGGACCAG GAGGAGCTAC TACCTCAATG AGATCCAGAG	540
CTTCGCGGGC GCCGAGAAGG ACGCGCGCGT GGTGGGCGAG ATCGCCTTCC AGCTGGACCG	600
CCGCATCCTG GCCTACGTGT TCCCGGGCGT GACGCGGCTC TACGGCTTCA CGGTGGCCAA	660
CATCCCCGAG AAGATCGAGC AGACCTCCAC CAAGTCTCTG GACGGCTCCG TGGACGAGAG	720
GAAGCTGCGC GAGCTGACGC AGCGCTACCT GGCCCTGAGC GCGCCCTGG AGAAGCTGGG	780
CTACAGCCGC GACGTGCACC CGGCGTTCAG CGAGTTCCTC ATCAACACCT ACGGAATCCT	840
GAAGCAGCGG CCCGACCTGC GCGCCAACCC CCTGCACAGC AGCCCGGCCG CGCTGCGCAA	900
GCTGGTCATC GACGTGGTGC CCCCCAAGTT CCTGGGCGAC TCGCTGCTGC TGCTCAACTG	960
CCTGTGCGAG CTCTCCAAGG AGGACGGCAA GCCCCTCTTC GCCTGGTGAG CCGCCCCGCG	1020
CCCGCCGCCT TGCCTGCAGT AAACGCGTTT GTTCCAACCC GGGGCCGCGG TGCTCCTGC	1080
GCGTCCCCCC GGAGGGGAAA GGGCCGCGTC CCCC GCGCGC GAGGCCAGAG AAGGCCCGC	1140
TCCCACCGGT GCTGGGCCCC GACCGCAGCC CGCCGCTGCC CGCACCTGCG GAGTGCTTCT	1200
CACCCCTCAT TAAATCATC CGTTTGCAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	1260
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA	1307

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 186 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

(2) INFORMATION FOR SEQ ID NO:19:

(A) LENGTH: 3319 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

86

CAGCATGAAA CACAAGCGCC TGTCCCGTCA TTCCACCGCC AGCCACAGCA GTTCCCACAC 180
CTCGGGCATT GAGGCAGACA CCAAGCCCCG GGACACGGGG CCGGAAGACA GCTACTCCAG 240
CAGTGCCATC CACCGCAAAGC TGAAAACCTG CAGCTCAATG ACCAGTCATG GCAGCTCCCA 300
CACCTCAGGG GTGGAGAGTG GCGGCAAAGA CCGGCTGGAA GAGGACTTAC AGGACGATGA 360
AATAGAGATG TTGGTTGATG ACCCCCGGGA TCTGGAGCAG ATGAATGAAG AGTCTCTGGA 420
AGTCAGCCCA GACATGTGCA TCTACATCAC AGAGGACATG CTCATGTCGC GGAAGCTGAA 480
TGGACACTCT GGGTTGATTG TGAAAGAAAT TGGGTCTTCC ACCTCGAGCT CTTCAGAAAC 540
AGTTGTTAAG CTTTCGTGGCC AGAGTACTGA TTCTCTTCCA CAGACTATAT GTCGGAAACC 600
AAAGACCTCC ACTGATCGAC ACAGCTTGAG CCTCGATGAC ATCAGACTTT ACCAGAAAGA 660
CTTCCTGCGC ATTGCAGGTC TGTGTCAGGA CACTGCTCAG AGTTACACCT TTGGATGTGG 720
CCATGAACTG GATGAGGAAG GCCTCTATTG CAACAGTTCC TTCCCCCACC AGTGCATCAA 780
CATCCAAGAT GCTTTTCCAG TCAAAGAAG CAGCAAATAC TTTTCTCTGG ATCTCACTCA 840
TGATGAAGTT CCAGAGTTTG TTGTGTAAAG TCCGTCTGTG TGCAGCTGTA CAGGCAGCTT 900
ACTGTTTGCT AGAGGATGCG AAAGTCATAA GTTCTTTACA TATTACTTGT GCCATATCTT 960
CTTCACCCTA AACATAGCTC TTTCTTTATA ATATTTGTGA TGATGGAAAC AAAAGCCTTG 1020
GAACAATTGC ACTTTAAGTA TTACACAGAA GTAAAAGAAC TACAGAAAAT GTACAGCAAG 1080
ACAAGTGCCC GGAAGTTCAC TGATCCTTCA GAAGGAAATG CGCTTTACTG ATTGCAAAGC 1140
CTTCAGAATA TTGGAGTGTG GTGTGTTTGC TCATCTGATG CTTTTTAGTT CAGTTACATG 1200
TAACATCACA TTTTTTTTAT CACGTGAAAG ATGTTAGATT TGTTTGCTTA TAAATTTTTT 1260
ACCACTCCCA CATAAAATGC TCATAGTTTG GGAGAGGAAA GAGGGAAGAT TCTCTCTTCT 1320
TTTAACAGAG AGATGATTGC TCTGTATACC CATTGCTTCC TCCCTGAGGC TGTCCCAAAG 1380
TGAACACTGA TGGAGTGGTC AAAATCATAA GATTGTAGCA AGCCAAAGAT ACGTATGTGA 1440
CGGAAGCACA TAAGCAATAA GCAGAAAACC AGAAGTGCAT GCTGTGATGC CTGTGACTCC 1500
TTCATCCCGC TCAGTGCCAT GTCCTCTTTT GTGATCTTCC AGAAAGCTCC AGGATTCATT 1560
TGAGTTCCAC ATCCAAGTAA CAGATGAATT ATATTCATGT TGTAATGCAT TTTGTGGAGT 1620
TTACAAAACC AGTGTCTGTT AAAACTTTGG AAAATGTCTT AGAAAACGTT GGTGCTTGGT 1680
GATGCTTTAT TTGTTTAATT ATCAAGAACA AATTATGGCA ATGCTAGTTT CTGCTTAACC 1740
AAAATACTCT GTGTATATAT TATACATATA TAAATACATG GGATTGTGTA TGTCTATATG 1800

TGTTTAAAGC TTACTATGTC TTCATTTTGG CTTCCATGAC TATCTTTTAT ACATGGAATT 1860
CCTTAAGATT GAGAATATGT CACTGAGTGA ATGATACCTG CAGACAGTCA GTTGATATAT 1920
GTAGAGTTCA GAATGACTGT TTTCTCATGT GCCTTTGGCC ATGATTCTCA AACTGATTG 1980
TATAACAGAA TTTTGGGGGG AGCTTTTAAA AAATAATGAC TGAGTCTCCC ACCAGACCGA 2040
TTACATCATT CTCTTGTTGGC GGGACCCAAG TAGAATTGCC TTTTCTTTTA AAGTTCTCCA 2100
GATGGAGCTA ATATGCAACA AAGTTGAAAA CCACTGATCC TGGGGGTGTC TTGTTAATTT 2160
TGAAGTAAAA GTGTACAGAA GACGTAGTGT ATGAGAAAGG GCCATTTTTA AGACAGTTAC 2220
CTGTTGTGCT GCTGTTACAA TATATAATGA AACCAAGTCA GGGGAGTGAA TTTATCAATC 2280
TTTTGATGTA AAGTAAAAAC GTAGTTCACA CTTCAGGAGA GAACTTCATA GCACAATGTC 2340
TTTCTATAAG ATATTTTAA TGATTTAGTA TTTTACAACA TTTGTTTACC ATATTTTGAT 2400
ATACCATTTT TTTCTATCTG CCCAGTTTTA TAAAAAAC TATATATTAT TTTCTAAAGA 2460
AACAAATCATA TTTTATACA AAATTATGTT TTCAGGTAAC GAAATAGATG TAGGGTACAG 2520
TGGAACATAA GCAGTGTTAC CCCTGGCTGG GAGTCAGTAT TATACAACAA ATGGTGAGCT 2580
GGAACATGCC CTGTCTGTGC TGTCCCTCCT GTGCTGGGTC GCGGATATGT AGGCAACATT 2640
GCCTTATCAC GCTAGGTTCA CCTGACACTT TAAAAGGAAA AAAAGTTCCA TAGAGTTCTG 2700
TGGTCACAAA ATTGTTTTGC TTTTATCAA TACTTTAATA GAACCAAAGT TGCAGATATT 2760
GGAATGTATG GAAGTATCTC AGTCTCTGCA TAAGAGGATT AAAGTATGAA AGGATCATTT 2820
AATGACTGTT TTAATTATAA GTCATTAAGT AATCCACCAT TTCTTATGGA TGATGCTTAA 2880
GCCTGGTGAG GTTTGTACTC TAAGGAGCCC AGATCATAAT GCAGTGCATT TCCTTAGCCC 2940
TTAGAGTTTC TTGCAACAT TTAACAAAAA GACATATTTA AGAAAGAAAG ATAAAGAAAA 3000
AACATATTTA ATTACTGTAA ACAGGTACTG CTTTATGTTT ATTTTCTCTC TACTTCAACC 3060
AAAATCAGAT CTTTGAGGTT TTGCTGACAT TGTTGGTGGT TTTGCACATG TTCTTTCTAA 3120
TTGGATTTAT GAATAGTTCT ATGGGTTTTT AAAGATGAAT CATGCTAAGA AACTTCTGTC 3180
TTTTTGATCC ACTGTTTGCA GCAGAATTAT ATATATGTAT AGGAAAAATC CACTTTGAAT 3240
AATCCATGTT TTGTATTTGG AAATTGTTTT TAAAAATAAA AAGGAAAGGA AATATAAAAA 3300
AAAAAAAAA AAAAAAAAAA 3319

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 264 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met	Asp	Gln	Leu	Glu	Lys	Arg	Ser	Arg	Ala	Ser	Gly	Ser	Ser	Ala	Gly	1	5	10	15
Ser	Met	Lys	His	Lys	Arg	Leu	Ser	Arg	His	Ser	Thr	Ala	Ser	His	Ser	20	25	30	
Ser	Ser	His	Thr	Ser	Gly	Ile	Glu	Ala	Asp	Thr	Lys	Pro	Arg	Asp	Thr	35	40	45	
Gly	Pro	Glu	Asp	Ser	Tyr	Ser	Ser	Ser	Ala	Ile	His	Arg	Lys	Leu	Lys	50	55	60	
Thr	Cys	Ser	Ser	Met	Thr	Ser	His	Gly	Ser	Ser	His	Thr	Ser	Gly	Val	65	70	75	80
Glu	Ser	Gly	Gly	Lys	Asp	Arg	Leu	Glu	Glu	Asp	Leu	Gln	Asp	Asp	Glu	85	90	95	
Ile	Glu	Met	Leu	Val	Asp	Asp	Pro	Arg	Asp	Leu	Glu	Gln	Met	Asn	Glu	100	105	110	
Glu	Ser	Leu	Glu	Val	Ser	Pro	Asp	Met	Cys	Ile	Tyr	Ile	Thr	Glu	Asp	115	120	125	
Met	Leu	Met	Ser	Arg	Lys	Leu	Asn	Gly	His	Ser	Gly	Leu	Ile	Val	Lys	130	135	140	
Glu	Ile	Gly	Ser	Ser	Thr	Ser	Ser	Ser	Ser	Glu	Thr	Val	Val	Lys	Leu	145	150	155	160
Arg	Gly	Gln	Ser	Thr	Asp	Ser	Leu	Pro	Gln	Thr	Ile	Cys	Arg	Lys	Pro	165	170	175	
Lys	Thr	Ser	Thr	Asp	Arg	His	Ser	Leu	Ser	Leu	Asp	Asp	Ile	Arg	Leu	180	185	190	
Tyr	Gln	Lys	Asp	Phe	Leu	Arg	Ile	Ala	Gly	Leu	Cys	Gln	Asp	Thr	Ala	195	200	205	
Gln	Ser	Tyr	Thr	Phe	Gly	Cys	Gly	His	Glu	Leu	Asp	Glu	Glu	Gly	Leu	210	215	220	
Tyr	Cys	Asn	Ser	Cys	Leu	Ala	Gln	Gln	Cys	Ile	Asn	Ile	Gln	Asp	Ala				

225

230

235

240

Phe Pro Val Lys Arg Thr Ser Lys Tyr Phe Ser Leu Asp Leu Thr His
245 250 255

Asp Glu Val Pro Glu Phe Val Val
260

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ANATTCAAGAG AGTTGAACTG AATAACCC

29

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ANGGCAGAGCC ACATCCTTCA CAACAGAA

29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GNGTCTGAGTG TCTATGTGAG GGCAAGGA

29

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TNGGATAAGCC GGCACAGACG AAGGCCAT

29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GNAAGGAAGGA GACAAACACA ATCACCCA

29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TNTTTGTTCTA GCAATTCTAC TTCCGGGA

29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TNCTTCCTCAC GTCTCTGTGC CAAATATT

29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ANCTCCAGTCT GATGCCAGTC ATGTCTAC

29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonulceotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GNCGGCAACAC TCGTGAACCT TCCGAAAT

29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GNCTTGGAACA ATTGCACTTT AAGTATTA

29

What is claimed is:

1. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 533 to nucleotide 673;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 596 to nucleotide 673;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 664;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bd379_1 deposited under accession number ATCC 98361;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bd379_1 deposited under accession number ATCC 98361;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bd379_1 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bd379_1 deposited under accession number ATCC 98361;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

2. A composition of claim 1 wherein said polynucleotide is operably linked to an expression control sequence.

3. A host cell transformed with a composition of claim 2.
4. The host cell of claim 3, wherein said cell is a mammalian cell.
5. A process for producing a protein, which comprises:
 - (a) growing a culture of the host cell of claim 3 in a suitable culture medium; and
 - (b) purifying the protein from the culture.
6. A protein produced according to the process of claim 5.
7. The protein of claim 6 comprising a mature protein.
8. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:
 - (a) the amino acid sequence of SEQ ID NO:2;
 - (b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44;
 - (c) fragments of the amino acid sequence of SEQ ID NO:2; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone bd379_1 deposited under accession number ATCC 98361;the protein being substantially free from other mammalian proteins.
9. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.
10. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44.
11. The composition of claim 8, further comprising a pharmaceutically acceptable carrier.

12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

14. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 55 to nucleotide 1008;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 952 to nucleotide 1008;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 403 to nucleotide 981;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bp121_2 deposited under accession number ATCC 98361;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bp121_2 deposited under accession number ATCC 98361;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bp121_2 deposited under accession number ATCC 98361;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bp121_2 deposited under accession number ATCC 98361;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

15. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
 - (b) the amino acid sequence of SEQ ID NO:4 from amino acid 119 to amino acid 309;
 - (c) fragments of the amino acid sequence of SEQ ID NO:4; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone bp121_2 deposited under accession number ATCC 98361;
- the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

17. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 52 to nucleotide 639;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 308;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bp646_10 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bp646_10 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bp646_10 deposited under accession number ATCC 98361;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bp646_10 deposited under accession number ATCC 98361;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

18. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 86;

(c) fragments of the amino acid sequence of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone bp646_10 deposited under accession number ATCC 98361;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

20. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 134 to nucleotide 1183;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 191 to nucleotide 1183;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 763;

- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cf50_1 deposited under accession number ATCC 98361;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cf50_1 deposited under accession number ATCC 98361;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cf50_1 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cf50_1 deposited under accession number ATCC 98361;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

21. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:8;
- (b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 210;
- (c) fragments of the amino acid sequence of SEQ ID NO:8; and
- (d) the amino acid sequence encoded by the cDNA insert of clone cf50_1 deposited under accession number ATCC 98361;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

23. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 740 to nucleotide 2245;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1 to nucleotide 463;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cw1543_3 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cw1543_3 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cw1543_3 deposited under accession number ATCC 98361;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cw1543_3 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:10;
- (b) fragments of the amino acid sequence of SEQ ID NO:10; and

(c) the amino acid sequence encoded by the cDNA insert of clone cw1543_3 deposited under accession number ATCC 98361; the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

26. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 952 to nucleotide 1074;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 524 to nucleotide 1059;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone da389_1 deposited under accession number ATCC 98361;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone da389_1 deposited under accession number ATCC 98361;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone da389_1 deposited under accession number ATCC 98361;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone da389_1 deposited under accession number ATCC 98361;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

27. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:12;
- (b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 36;
- (c) fragments of the amino acid sequence of SEQ ID NO:12; and
- (d) the amino acid sequence encoded by the cDNA insert of clone da389_1 deposited under accession number ATCC 98361;

the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11.

29. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 78 to nucleotide 1619;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 604 to nucleotide 1307;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dd71_2 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dd71_2 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone dd71_2 deposited under accession number ATCC 98361;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone dd71_2 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity;

- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

30. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- (b) the amino acid sequence of SEQ ID NO:14 from amino acid 200 to amino acid 410;
- (c) fragments of the amino acid sequence of SEQ ID NO:14; and
- (d) the amino acid sequence encoded by the cDNA insert of clone dd71_2 deposited under accession number ATCC 98361;

the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.

32. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1003 to nucleotide 1350;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 714 to nucleotide 1320;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dm221_1 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dm221_1 deposited under accession number ATCC 98361;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone dm221_1 deposited under accession number ATCC 98361;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone dm221_1 deposited under accession number ATCC 98361;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

33. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 106;

(c) fragments of the amino acid sequence of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone dm221_1 deposited under accession number ATCC 98361;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

35. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 449 to nucleotide 1006;

- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 1 to nucleotide 331;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dx279_1 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dx279_1 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone dx279_1 deposited under accession number ATCC 98361;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone dx279_1 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

36. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
 - (b) fragments of the amino acid sequence of SEQ ID NO:18; and
 - (c) the amino acid sequence encoded by the cDNA insert of clone dx279_1 deposited under accession number ATCC 98361;
- the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17.

38. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 74 to nucleotide 865;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 538 to nucleotide 1044;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gm243_1 deposited under accession number ATCC 98361;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gm243_1 deposited under accession number ATCC 98361;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone gm243_1 deposited under accession number ATCC 98361;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone gm243_1 deposited under accession number ATCC 98361;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

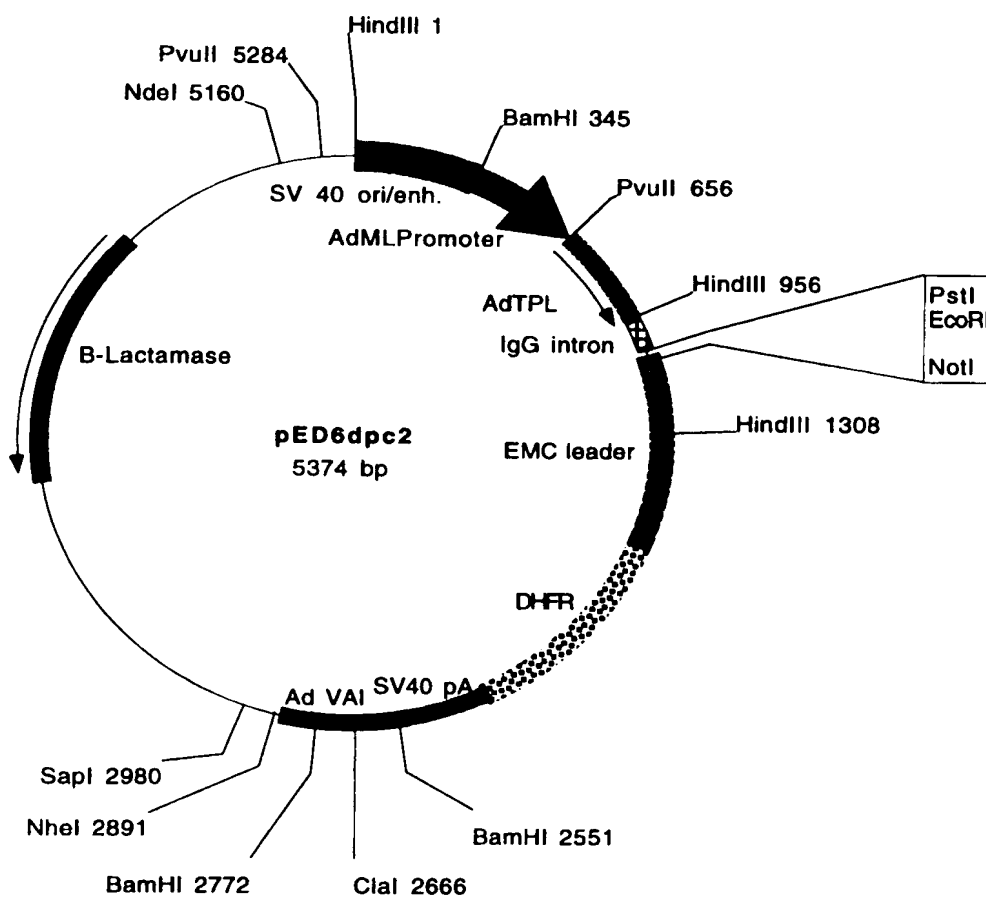
39. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:20;
- (b) fragments of the amino acid sequence of SEQ ID NO:20; and

(c) the amino acid sequence encoded by the cDNA insert of clone gm243_1 deposited under accession number ATCC 98361; the protein being substantially free from other mammalian proteins.

40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:19.

FIGURE 1A

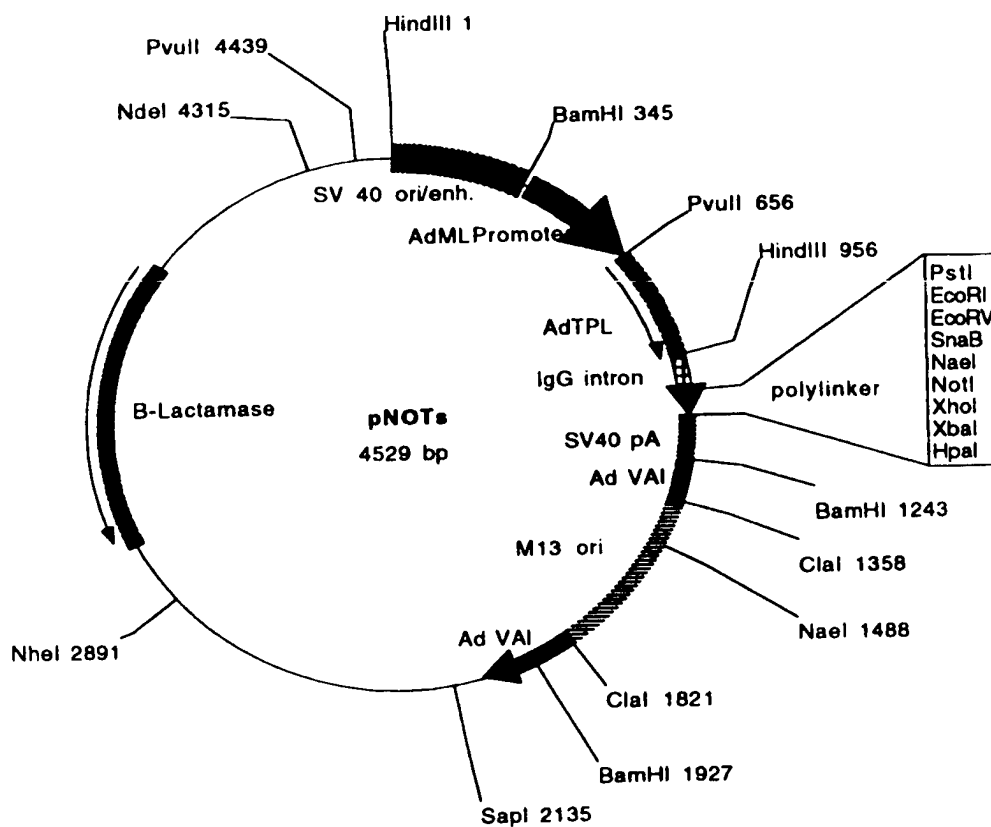


Plasmid name: pED6dpc2

Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs

Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al, 1989. Mol. Cell. Biol. 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the Clal site. SST cDNAs are cloned between EcoRI and NotI



Comments/References: pED6pc2 is derived from pED6pc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991). NAR 19: 4485-4490.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

Intern 1al Application No

PCT/US 98/04977

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N5/10 C07K14/47 C12Q1/68 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K C12Q A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	C. AUFRAY ET AL.: "The Genexpress cDNA program" EMBL SEQUENCE DATABASE, 12 February 1995, HEIDELBERG, FRG, XP002069658 cited in the application H. sapiens partial cDNA sequence; clone c-05b06; Accession no. F05256; ---	1,13
X	L. HILLIER ET AL.: "The WashU-Merck EST Project" EMBL SEQUENCE DATABASE, 29 May 1995, HEIDELBERG, FRG, XP002069659 cited in the application yh04b03.r1 Homo sapiens cDNA clone 42053 5'; Accession no. R60369; --- -/-	1,13

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

° Special categories of cited documents :

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

Z document member of the same patent family

Date of the actual completion of the international search

29 June 1998

Date of mailing of the international search report

30.09.98

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

HORNIG H.

INTERNATIONAL SEARCH REPORT

Internat Application No

PCT/US 98/04977

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	L. HILLIER ET AL.: "The WashU-Merck EST Project" EMBL SEQUENCE DATABASE, 17 May 1996, HEIDELBERG, FRG, XP002069660 cited in the application zcl8g02.r1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 322706 5'; Accession no. W39550; ---	1,13
A	ADAMS M D ET AL: "3,400 NEW EXPRESSED SEQUENCE TAGS IDENTIFY DIVERSITY OF TRANSCRIPTS IN HUMAN BRAIN" NATURE GENETICS, vol. 4, no. 3, pages 256-267, XP000611495 see the whole document ---	1-13
A	JACOBS K ET AL: "A NOVEL METHOD FOR ISOLATING EUKARYOTIC CDNA CLONES ENCODING SECRETED PROTEINS" JOURNAL OF CELLULAR BIOCHEMISTRY - SUPPLEMENT, vol. 21A, 10 March 1995, page 19 XP002027246 see abstract ---	1-13
A	EP 0 510 691 A (OSAKA BIOSCIENCE INST) 28 October 1992 see the whole document ---	1-13
A	WO 94 07916 A (MERCK & CO INC ;SCHMIDT AZRIEL (US); RODAN GIDEON A (US); RUTLEDGE) 14 April 1994 see the whole document ---	1-13
A	WO 90 05780 A (OREGON STATE) 31 May 1990 see the whole document ---	1-13
A	WO 90 14432 A (GENETICS INST) 29 November 1990 see the whole document ---	1-13
A	WO 96 17925 A (IMMUNEX CORP) 13 June 1996 see the whole document ---	1-13
A	R.J. KAUFMAN ET AL.: "Effect of von Willebrand factor coexpression on the synthesis and secretion of factor VIII in chinese hamster ovary cells" MOL. CELL. BIOL., vol. 9, no. 3, March 1989, ASM WASHINGTON, DC,US, pages 1233-1242, XP002041592 see the whole document ---	1-13

-/--

INTERNATIONAL SEARCH REPORT

Intern: ional Application No

PCT/JP98/04977

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	R.J. KAUFMAN ET AL.: "The phosphorylation state of eucaryotic initiation factor 2 alters translation efficiency of specific mRNAs" MOL. CELL. BIOL., vol. 9, no. 3, March 1989, ASM WASHINGTON, DC,US, pages 946-958, XP002041593 see the whole document ---	1-13
A	R.J. KAUFMAN ET AL.: "Improved vectors for stable expression of foreign genes in mammalian cells by use of the untranslated leader sequence from EMC virus" NUCLEIC ACIDS RESEARCH, vol. 19, no. 16, 1991, IRL PRESS LIMITED,OXFORD,ENGLAND, pages 4485-4490, XP002041594 cited in the application see the whole document ---	1-13
A	US 5 536 637 A (JACOBS KENNETH) 16 July 1996 cited in the application see the whole document ---	1-13
A	WO 97 07198 A (GENETICS INSTITUT) 27 February 1997 see the whole document ---	1-13
P,A	WO 97 25427 A (GENETICS INST) 17 July 1997 see the whole document -----	1-13

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 98/ 04977

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim 12 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see further information sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-13

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-13

A composition comprising an isolated polynucleotide selected from the group consisting of: SEQ ID no.1; said composition wherein said polynucleotide is operably linked to an expression control sequence; a host cell transformed with said composition; a process for producing a protein which is encoded by said polynucleotide sequence; a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group of SEQ ID no.2, said composition further comprising a pharmaceutical acceptable carrier; a method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of said composition, the gene corresponding to the cDNA sequence of SEQ ID no.1.

2. Claims: 14-16

A composition comprising an isolated polynucleotide sequence selected from the group of SEQ ID no.3; a composition comprises a protein, wherein said protein comprises an amino acid sequence selected from the group of SEQ ID no.4; the gene corresponding to the cDNA sequences of SEQ ID no.3;

3. Claims: 17-19

Idem as subject 2 but limited to SEQ ID nos.5 and 6.

4. Claims: 20-22

Idem as subject 2 but limited to SEQ ID nos.7 and 8.

5. Claims: 23-25

Idem as subject 2 but limited to SEQ ID nos.9 and 10.

6. Claims: 26-28

Idem as subject 2 but limited to SEQ ID nos.11 and 12.

7. Claims: 29-31

Idem as subject 2 but limited to SEQ ID nos.13 and 14.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

8. Claims: 32-34

Idem as subject 2 but limited to SEQ ID nos.15 and 16.

9. Claims: 35-37

Idem as subject 2 but limited to SEQ ID nos.17 and 18.

10. Claims: 38-40

Idem as subject 2 but limited to SEQ ID nos.19 and 20.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/JP98/04977

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0510691 A	28-10-92	CA 2067031 A JP 5184368 A	27-10-92 27-07-93
WO 9407916 A	14-04-94	AU 5165193 A	26-04-94
WO 9005780 A	31-05-90	AT 154636 T AU 645963 B AU 4668689 A DE 68928137 D DE 68928137 T EP 0447483 A ES 2104599 T JP 4506449 T	15-07-97 03-02-94 12-06-90 24-07-97 19-02-98 25-09-91 16-10-97 12-11-92
WO 9014432 A	29-11-90	US 5580753 A AT 147436 T AU 637620 B AU 5928990 A CA 2056997 A DE 69029657 D DK 473724 T EP 0473724 A ES 2099096 T JP 4506006 T US 5734037 A US 5414071 A	03-12-96 15-01-97 03-06-93 18-12-90 24-11-90 20-02-97 14-04-97 11-03-92 16-05-97 22-10-92 31-03-98 09-05-95
WO 9617925 A	13-06-96	AU 4639396 A CA 2206488 A FI 972390 A NO 972455 A	26-06-96 13-06-96 05-06-97 06-08-97
US 5536637 A	16-07-96	US 5712116 A	27-01-98
WO 9707198 A	27-02-97	US 5707829 A AU 6712396 A AU 6768596 A EP 0839196 A EP 0851875 A WO 9704097 A	13-01-98 18-02-97 12-03-97 06-05-98 08-07-98 06-02-97

Information on patent family members

PCT/US 98/04977

01-08-97



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, C07K 14/47, A61K 38/17		A2		(11) International Publication Number: WO 98/32853
				(43) International Publication Date: 30 July 1998 (30.07.98)

<p>(21) International Application Number: PCT/US98/01396</p> <p>(22) International Filing Date: 23 January 1998 (23.01.98)</p> <p>(30) Priority Data: 08/788,789 24 January 1997 (24.01.97) US</p> <p>(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).</p> <p>(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).</p> <p>(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).</p>	<p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Published Without international search report and to be republished upon receipt of that report.</p>
--	--

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.

Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

5

10

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

15

This application is a continuation-in-part of Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/788,789), filed January 24, 1997, which is incorporated by reference herein.

20

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

25

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

40

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 5 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 506 to nucleotide 643;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 471 to nucleotide 765;
- 10 (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AA35_2 deposited under accession number ATCC 98303;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AA35_2 deposited under accession number ATCC 98303;
- 15 (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AA35_2 deposited under accession number ATCC 98303;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AA35_2 deposited under accession number ATCC 98303;
- 20 (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of
- 25 (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

30 Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 506 to nucleotide 643; the nucleotide sequence of SEQ ID NO:1 from nucleotide 471 to nucleotide 765; the nucleotide sequence of the full-length protein coding sequence of clone AA35_2 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone AA35_2 deposited

under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AA35_2 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein
5 comprising the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group
10 consisting of:

- (a) the amino acid sequence of SEQ ID NO:2;
- (b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32;
- (c) fragments of the amino acid sequence of SEQ ID NO:2; and
- 15 (d) the amino acid sequence encoded by the cDNA insert of clone AA35_2 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32.

20 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID
25 NO:3 from nucleotide 71 to nucleotide 736;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 113 to nucleotide 736;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 343;
- 30 (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AM42_3 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AM42_3 deposited under accession number ATCC 98303;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AM42_3 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AM42_3 deposited under accession number ATCC 98303;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 71 to nucleotide 736; the nucleotide sequence of SEQ ID NO:3 from nucleotide 113 to nucleotide 736; the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 343; the nucleotide sequence of the full-length protein coding sequence of clone AM42_3 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone AM42_3 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AM42_3 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91;

- (c) fragments of the amino acid sequence of SEQ ID NO:4; and
- (d) the amino acid sequence encoded by the cDNA insert of clone

AM42_3 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such

- 5 protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 10 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 55 to nucleotide 423;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG137_7 deposited under accession
15 number ATCC 98303;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG137_7 deposited under accession number ATCC 98303;
- (e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG137_7 deposited under accession number
20 ATCC 98303;
- (f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG137_7 deposited under accession number ATCC 98303;
- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;
- 25 (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity;
- (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;
- (j) a polynucleotide which encodes a species homologue of the protein
30 of (g) or (h) above ; and
- (k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 55 to nucleotide 423; the nucleotide sequence of the full-length

protein coding sequence of clone BG137_7 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone BG137_7 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BG137_7 deposited under accession number ATCC 98303. In yet other preferred
5 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 62 to amino acid 123.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ
10 ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:6;
- 15 (b) the amino acid sequence of SEQ ID NO:6 from amino acid 62 to amino acid 123;
- (c) fragments of the amino acid sequence of SEQ ID NO:6; and
- (d) the amino acid sequence encoded by the cDNA insert of clone BG137_7 deposited under accession number ATCC 98303;

20 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 62 to amino acid 123.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 25 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 186 to nucleotide 2030;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID
30 NO:7 from nucleotide 873 to nucleotide 2030;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 802 to nucleotide 1173;

- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CH699_1 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CH699_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CH699_1 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CH699_1 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 186 to nucleotide 2030; the nucleotide sequence of SEQ ID NO:7 from nucleotide 873 to nucleotide 2030; the nucleotide sequence of SEQ ID NO:7 from nucleotide 802 to nucleotide 1173; the nucleotide sequence of the full-length protein coding sequence of clone CH699_1 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone CH699_1 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CH699_1 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8 from amino acid 218 to amino acid 329.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:8;
- 5 (b) the amino acid sequence of SEQ ID NO:8 from amino acid 218 to amino acid 329;
- (c) fragments of the amino acid sequence of SEQ ID NO:8; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CH699_1 deposited under accession number ATCC 98303;
- 10 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence of SEQ ID NO:8 from amino acid 218 to amino acid 329.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 15 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 111 to nucleotide 677;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID
- 20 NO:10 from nucleotide 156 to nucleotide 677;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO851_1 deposited under accession number ATCC 98303;
- (e) a polynucleotide encoding the full-length protein encoded by the
- 25 cDNA insert of clone CO851_1 deposited under accession number ATCC 98303;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO851_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide encoding the mature protein encoded by the
- 30 cDNA insert of clone CO851_1 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

5 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:10 from nucleotide 111 to nucleotide 677; the nucleotide sequence of SEQ ID NO:10 from nucleotide 156 to nucleotide 677; the nucleotide sequence of the full-length protein coding sequence of clone CO851_1 deposited under accession number ATCC 98303; or the
10 nucleotide sequence of the mature protein coding sequence of clone CO851_1 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO851_1 deposited under accession number ATCC 98303. In yet other preferred
15 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11 from amino acid 120 to amino acid 189.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:10, SEQ ID NO:9 or SEQ ID NO:12 .

20 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:11;

(b) the amino acid sequence of SEQ ID NO:11 from amino acid 120 to
25 amino acid 189;

(c) fragments of the amino acid sequence of SEQ ID NO:11; and

(d) the amino acid sequence encoded by the cDNA insert of clone
CO851_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such
30 protein comprises the amino acid sequence of SEQ ID NO:11 or the amino acid sequence of SEQ ID NO:11 from amino acid 120 to amino acid 189.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 123 to nucleotide 755;

5 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 279 to nucleotide 755;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 631;

10 (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CP111_1 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CP111_1 deposited under accession number ATCC 98303;

15 (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CP111_1 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CP111_1 deposited under accession number ATCC 98303;

20 (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

25 (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

30 Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 123 to nucleotide 755; the nucleotide sequence of SEQ ID NO:13 from nucleotide 279 to nucleotide 755; the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 631; the nucleotide sequence of the full-length protein coding sequence of clone CP111_1 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone CP111_1 deposited

under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CP111_1 deposited under accession number ATCC 98303. In yet other preferred
5 comprising the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 171.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising
10 a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- (b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 171;
- 15 (c) fragments of the amino acid sequence of SEQ ID NO:14; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CP111_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14 or the amino acid sequence
20 of SEQ ID NO:14 from amino acid 1 to amino acid 171.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- 25 (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 214 to nucleotide 2760;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 406 to nucleotide 2760;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID
30 NO:15 from nucleotide 2011 to nucleotide 2565;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CS278_1 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CS278_1 deposited under accession number ATCC 98303;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CS278_1 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CS278_1 deposited under accession number ATCC 98303;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 214 to nucleotide 2760; the nucleotide sequence of SEQ ID NO:15 from nucleotide 406 to nucleotide 2760; the nucleotide sequence of SEQ ID NO:15 from nucleotide 2011 to nucleotide 2565; the nucleotide sequence of the full-length protein coding sequence of clone CS278_1 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone CS278_1 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CS278_1 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 596 to amino acid 784.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:16;
- (b) the amino acid sequence of SEQ ID NO:16 from amino acid 596 to amino acid 784;

(c) fragments of the amino acid sequence of SEQ ID NO:16; and

- 5 (d) the amino acid sequence encoded by the cDNA insert of clone CS278_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 596 to amino acid 784.

- 10 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

- 15 (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 901 to nucleotide 1074;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 970 to nucleotide 1074;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 626 to nucleotide 1147;

- 20 (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DF968_3 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DF968_3 deposited under accession number ATCC 98303;

- 25 (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DF968_3 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DF968_3 deposited under accession number ATCC 98303;

- 30 (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

5 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 901 to nucleotide 1074; the nucleotide sequence of SEQ ID NO:17 from nucleotide 970 to nucleotide 1074; the nucleotide sequence of SEQ ID NO:17 from nucleotide 626 to nucleotide 1147; the nucleotide sequence of the full-length protein coding sequence of clone DF968_3 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone DF968_3 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DF968_3 deposited under accession number ATCC 98303.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
- (b) fragments of the amino acid sequence of SEQ ID NO:18; and
- (c) the amino acid sequence encoded by the cDNA insert of clone DF968_3 deposited under accession number ATCC 98303;

25 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 560 to nucleotide 820;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DN1120_2 deposited under accession number ATCC 98303;

5 (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DN1120_2 deposited under accession number ATCC 98303;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DN1120_2 deposited under accession number ATCC 98303;

10 (f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DN1120_2 deposited under accession number ATCC 98303;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity;

15 (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

20 (k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 560 to nucleotide 820; the nucleotide sequence of the full-length protein coding sequence of clone DN1120_2 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone
25 DN1120_2 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DN1120_2 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20 from amino acid
30 1 to amino acid 61.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:20;
- 5 (b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 61;
- (c) fragments of the amino acid sequence of SEQ ID NO:20; and
- (d) the amino acid sequence encoded by the cDNA insert of clone DN1120_2 deposited under accession number ATCC 98303;
- 10 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:20 or the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 61.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 15 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 649 to nucleotide 786;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID
- 20 NO:21 from nucleotide 736 to nucleotide 786;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 525 to nucleotide 787;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DO589_1 deposited under accession
- 25 number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DO589_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DO589_1 deposited under accession number
- 30 ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DO589_1 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:22;

- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:22 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- 5 (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:21 from nucleotide 649 to nucleotide 786; the nucleotide sequence of SEQ ID NO:21 from nucleotide 736 to nucleotide 786; the nucleotide sequence of SEQ ID NO:21 from nucleotide 525 to nucleotide 787; the nucleotide sequence of the full-length protein coding sequence of clone DO589_1 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone DO589_1 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DO589_1 deposited under accession number ATCC 98303.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:21.

20 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:22;
- (b) fragments of the amino acid sequence of SEQ ID NO:22; and
- 25 (c) the amino acid sequence encoded by the cDNA insert of clone DO589_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:22.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

- (a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and
- (b) purifying the protein from the culture.

5 The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which
10 specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

15

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

20

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone
25 in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have
30 determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation

proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

5 Clone "AA35_2"

A polynucleotide of the present invention has been identified as clone "AA35_2". AA35_2 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer
10 analysis of the amino acid sequence of the encoded protein. AA35_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AA35_2 protein").

The nucleotide sequence of AA35_2 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the
15 predicted amino acid sequence of the AA35_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AA35_2 should be approximately 1400 bp.

The nucleotide sequence disclosed herein for AA35_2 was searched against the
20 GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AA35_2 demonstrated at least some similarity with sequences identified as C16789 (Human placenta cDNA 5'-end GEN-529D11), H23653 (yn72e01.r1 Homo sapiens cDNA clone 173976 5' similar to contains Alu repetitive element), L31848 (Homo sapiens serine/threonine kinase receptor 2 (SKR2) gene, 3 alternative splices, 3'
25 ends), U40455 (Human chromosome X cosmid, clones 196B12, 9H11 and 43H9, repeat units and sequence tagged sites), and Z82197 (Human DNA sequence from clone J293L6). The predicted amino acid sequence disclosed herein for AA35_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted AA35_2 protein demonstrated at least some similarity to sequences
30 identified as U58658 (unknown [Homo sapiens]) and X55777 (put. ORF [Homo sapiens]). Based upon sequence similarity, AA35_2 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of AA35_2 indicates that it may contain an Alu repetitive element.

Clone "AM42_3"

A polynucleotide of the present invention has been identified as clone "AM42_3". AM42_3 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was
5 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. AM42_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AM42_3 protein").

The nucleotide sequence of AM42_3 as presently determined is reported in SEQ
10 ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the AM42_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 2 to 14 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 15, or are a transmembrane domain.

15 The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AM42_3 should be approximately 1400 bp.

The nucleotide sequence disclosed herein for AM42_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AM42_3 demonstrated at least some similarity with sequences
20 identified as AA109637 (mm01f02.r1 Stratagene mouse kidney (#937315) Mus musculus cDNA clone 520251 5'), AA131170 (zo08e05.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone 567104 3'), AA131483 (zo08e05.r1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone 567104 5'), and AA445683 (vf62h07.r1 Barstead MPLRB1 Mus musculus cDNA clone 848413 5'). Based upon
25 sequence similarity, AM42_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the AM42_3 protein sequence centered around amino acid 152 of SEQ ID NO:4.

Clone "BG137_7"

30 A polynucleotide of the present invention has been identified as clone "BG137_7". BG137_7 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer

analysis of the amino acid sequence of the encoded protein. BG137_7 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BG137_7 protein").

5 The nucleotide sequence of BG137_7 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the BG137_7 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BG137_7 should be approximately 500 bp.

10 The nucleotide sequence disclosed herein for BG137_7 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BG137_7 demonstrated at least some similarity with sequences identified as D87683 (Human mRNA for KIAA0243 gene, partial cds). Based upon sequence similarity, BG137_7 proteins and each similar protein or peptide may share at
15 least some activity.

Clone "CH699_1"

A polynucleotide of the present invention has been identified as clone "CH699_1". CH699_1 was isolated from a human fetal kidney cDNA library using methods which are
20 selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CH699_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CH699_1 protein").

25 The nucleotide sequence of CH699_1 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CH699_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 217 to 229 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at
30 amino acid 230, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CH699_1 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for CH699_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and

FASTA search protocols. CH699_1 demonstrated at least some similarity with sequences identified as AA155014 (mr99h05.r1 Stratagene mouse embryonic carcinoma (#937317) Mus musculus cDNA clone 605625 5'), AA423476 (ve76d07.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone 832141 5'), U79271 (Human clones 23920 and 5 23921 mRNA sequence), and W72147 (zd70f08.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 346023 3'). The predicted amino acid sequence disclosed herein for CH699_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CH699_1 protein demonstrated at least some similarity to sequences identified as X51591 (beta-myosin heavy chain [Homo 10 sapiens]). Based upon sequence similarity, CH699_1 proteins and each similar protein or peptide may share at least some activity.

Clone "CO851_1"

A polynucleotide of the present invention has been identified as clone "CO851_1". 15 CO851_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CO851_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein 20 as "CO851_1 protein").

The nucleotide sequence of the 5' portion of CO851_1 as presently determined is reported in SEQ ID NO:9. An additional internal nucleotide sequence from CO851_1 as presently determined is reported in SEQ ID NO:10. What applicants believe is the proper reading frame and the predicted amino acid sequence encoded by such internal sequence 25 is reported in SEQ ID NO:11. Amino acids 3 to 15 of SEQ ID NO:11 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 16, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of CO851_1, including the polyA tail, is reported in SEQ ID NO:12.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone 30 CO851_1 should be approximately 1800 bp.

The nucleotide sequence disclosed herein for CO851_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CO851_1 demonstrated at least some similarity with sequences identified as AA132585 (zo20c04.r1 Stratagene colon (#937204) Homo sapiens cDNA clone

587430 5'), H51262 (yp83b07.s1 Homo sapiens cDNA clone 194005 3'), W44070 (mc73a09.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone 354136 5'), and X92871 (X.laevis mRNA for an unknown transmembrane protein). The predicted amino acid sequence disclosed herein for CO851_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CO851_1 protein demonstrated at least some similarity to sequences identified as X92871 (unknown transmembrane protein [Xenopus laevis]). Based upon sequence similarity, CO851_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of CO851_1 indicates that it may contain an Alu repetitive element.

Clone "CP111_1"

A polynucleotide of the present invention has been identified as clone "CP111_1". CP111_1 was isolated from a human adult salivary gland cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CP111_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CP111_1 protein").

The nucleotide sequence of CP111_1 as presently determined is reported in SEQ ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CP111_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14. Amino acids 40 to 52 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 53, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CP111_1 should be approximately 3200 bp.

The nucleotide sequence disclosed herein for CP111_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CP111_1 demonstrated at least some similarity with sequences identified as T53688 (ya98g07.r1 Homo sapiens cDNA clone 69756 5') and W70295 (zd58f03.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 344861 3'). The predicted amino acid sequence disclosed herein for CP111_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol.

The predicted CP111_1 protein demonstrated at least some similarity to sequences identified as X88852 (env protein [Primate T-cell lymphotropic]). Based upon sequence similarity, CP111_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the CP111_1 protein sequence centered around amino acid 50 of SEQ ID NO:14.

Clone "CS278_1"

A polynucleotide of the present invention has been identified as clone "CS278_1". CS278_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CS278_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CS278_1 protein").

The nucleotide sequence of CS278_1 as presently determined is reported in SEQ ID NO:15. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CS278_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16. Amino acids 52 to 64 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 65, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CS278_1 should be approximately 4400 bp.

The nucleotide sequence disclosed herein for CS278_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CS278_1 demonstrated at least some similarity with sequences identified as AA234319 (zr66c07.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 668364 5'), H44192 (yo73f09.r1 Homo sapiens cDNA clone 183593 5'), W18258 (mb86a11.r1 Soares mouse p3NMF19), X76589 (H.sapiens DNA 3' flanking simple sequence region clone wg2c3), and Z74652 (M.musculus mRNA; expressed sequence tag (tcc2)). The predicted amino acid sequence disclosed herein for CS278_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CS278_1 protein demonstrated at least some similarity to sequences identified as M34651 (ORF-3 protein [Suid herpesvirus 1]). The predicted CS278_1 protein also demonstrated at least some similarity to a protein motif, cytochrome P450 cysteine heme-

iron ligand signature. Based upon sequence similarity, CS278_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts five potential transmembrane domains within the CS278_1 protein sequence, which are centered around amino acids 75, 160, 525, 610, and 700 of SEQ ID NO:16, respectively. The nucleotide sequence of CS278_1 may contain GAA simple repeat elements.

Clone "DF968_3"

A polynucleotide of the present invention has been identified as clone "DF968_3". DF968_3 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DF968_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "DF968_3 protein").

The nucleotide sequence of DF968_3 as presently determined is reported in SEQ ID NO:17. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DF968_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18. Amino acids 11 to 23 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 24, or are a transmembrane domain. Another possible DF968_3 reading frame and predicted amino acid sequence is encoded by basepairs 191 to 430 of SEQ ID NO:17 and is reported in SEQ ID NO:33.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DF968_3 should be approximately 1010 bp.

The nucleotide sequence disclosed herein for DF968_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DF968_3 demonstrated at least some similarity with sequences identified as AA426010 (zw49e12.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 773422 3' similar to contains element LTR5 repetitive element), H18256 (yn48a04.r1 Homo sapiens cDNA clone 171630 5'), and T06820 (EST04709 Homo sapiens cDNA clone HFBDZ29). The predicted amino acid sequence disclosed herein for DF968_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted DF968_3 protein demonstrated at least some

similarity to sequences identified as Z38125 (orf, len 112, CAI 0.07). Based upon sequence similarity, DF968_3 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of DF968_3 indicates that it may contain repeat sequences.

5

Clone "DN1120_2"

A polynucleotide of the present invention has been identified as clone "DN1120_2". DN1120_2 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was
10 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DN1120_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "DN1120_2 protein").

The nucleotide sequence of DN1120_2 as presently determined is reported in SEQ
15 ID NO:19. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DN1120_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DN1120_2 should be approximately 1000 bp.

20 The nucleotide sequence disclosed herein for DN1120_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DN1120_2 demonstrated at least some similarity with sequences identified as M62256 (EST00323 Homo sapiens cDNA clone HHCH15 similar to Alu repetitive element), M78991 (EST01139 Homo sapiens cDNA clone HHCPG39), Q59179
25 (Human brain Expressed Sequence Tag EST00323), and Q61084 (Human brain Expressed Sequence Tag EST01139). Based upon sequence similarity, DN1120_2 proteins and each similar protein or peptide may share at least some activity.

Clone "DO589_1"

30 A polynucleotide of the present invention has been identified as clone "DO589_1". DO589_1 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DO589_1 is a full-length

clone, including the entire coding sequence of a secreted protein (also referred to herein as "DO589_1 protein").

The nucleotide sequence of DO589_1 as presently determined is reported in SEQ ID NO:21. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DO589_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:22. Amino acids 17 to 29 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 30, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DO589_1 should be approximately 1800 bp.

The nucleotide sequence disclosed herein for DO589_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DO589_1 demonstrated at least some similarity with sequences identified as AA402420 (zu47e04.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 741150 3'), AA426621 (zw03a09.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 768184 5'), AA436749 (zv67c10.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 758706 5'), H12845 (yj14h06.r1 Homo sapiens cDNA clone 148763 5'), R42350 (yg01b05.s1 Homo sapiens cDNA clone 30909 3'), W02775 (zc65g07.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 327228 3'), W24833 (zc65g07.r1 Soares fetal heart), W58173 (zd19f02.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 341115 3') similar to contains Alu repetitive element; contains element L1 repetitive element), and Z82201 (Human DNA sequence from clone J345P10). Based upon sequence similarity, DO589_1 proteins and each similar protein or peptide may share at least some activity.

Deposit of Clones

Clones AA35_2, AM42_3, BG137_7, CH699_1, CO851_1, CP111_1, CS278_1, DF968_3, DN1120_2, and DO589_1 were deposited on January 23, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98303, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited

by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* **9**: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

<u>Clone</u>	<u>Probe Sequence</u>
AA35_2	SEQ ID NO:23
AM42_3	SEQ ID NO:24
BG137_7	SEQ ID NO:25
CH699_1	SEQ ID NO:26
CO851_1	SEQ ID NO:27
CP111_1	SEQ ID NO:28
CS278_1	SEQ ID NO:29
DF968_3	SEQ ID NO:30
DN1120_2	SEQ ID NO:31
DO589_1	SEQ ID NO:32

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

- (a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;
- (b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with γ - ^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately $4\text{e}+6$ dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 μl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 $\mu\text{g}/\text{ml}$. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 $\mu\text{g}/\text{ml}$ and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 $\mu\text{g}/\text{ml}$ of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to $1\text{e}+6$ dpm/mL. The filter is then preferably

incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The
5 filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis,
10 hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S.
15 McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion
20 could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence
25 listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

30 The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited

to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed
5 sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s)
10 corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15**(7): 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62**(1): 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-
15 39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce
20 gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through
25 deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14**(9): 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90**(16): 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91**(2): 719-722; all of which are incorporated by reference herein), or through homologous recombination,
30 preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614,396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for

the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) [†]	Hybridization Temperature and Buffer [†]	Wash Temperature and Buffer [†]
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T _B *; 1xSSC	T _B *; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T _D *; 1xSSC	T _D *; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T _F *; 1xSSC	T _F *; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T _H *; 4xSSC	T _H *; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T _J *; 4xSSC	T _J *; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T _L *; 2xSSC	T _L *; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T _N *; 6xSSC	T _N *; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T _P *; 6xSSC	T _P *; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T _R *; 4xSSC	T _R *; 4xSSC

[†]: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed

to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

5 *: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH₂PO₄, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

10 * $T_B - T_R$: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, $T_m(^{\circ}\text{C}) = 2(\# \text{ of A} + \text{T bases}) + 4(\# \text{ of G} + \text{C bases})$. For hybrids between 18 and 49 base pairs in length, $T_m(^{\circ}\text{C}) = 81.5 + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\% \text{G} + \text{C}) - (600/\text{N})$, where N is the number of bases in the hybrid, and $[\text{Na}^+]$ is the concentration of sodium ions in the hybridization buffer ($[\text{Na}^+]$ for 1xSSC = 0.165 M).

15 Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

20 Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing
25 polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

30 The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control
35 sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205

cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as

those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and
5 subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or
10 all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic
15 animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are
20 known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic
25 compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA
30 sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art

(see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

10 The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies
15 or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express
20 recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare
25 with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for
30 examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that

described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

5 The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

20 Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

25 Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bertagnolli et al., *J. Immunol.* 145:1706-1712, 1990; Bertagnolli et al., *Cellular Immunology* 133:327-341, 1991; Bertagnolli, et al., *J. Immunol.* 149:3778-3783, 1992; Bowman et al., *J. Immunol.* 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ , Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., *J. Exp. Med.* 173:1205-1211, 1991; Moreau et al., *Nature* 336:690-692, 1988; Greenberger et al., *Proc. Natl. Acad. Sci. U.S.A.* 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in*

Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; 5 Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: 10 *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., 15 *Proc. Natl. Acad. Sci. USA* 77:6091-6095, 1980; Weinberger et al., *Eur. J. Immun.* 11:405-411, 1981; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

20 A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well 25 as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, 30 herpesviruses, mycobacteria, *Leishmania* spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the

molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to
5 anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

10 The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as
15 described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

20 Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms.
25 Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from
30 the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and

murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy.

- 5 Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B
- 10 lymphocyte antigens systemically.

- Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in*
- 15 *vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a
- 20 costimulatory signal to, and thereby activate, T cells *in vivo*.

- In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present
- 25 invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The
- 30 transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary

costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowman et al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro*

antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

5 Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988; Bertagnolli et al., *J. Immunol.* 149:3778-3783, 1992.

10 Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., *J. Immunol.* 134:536-544, 1995; Inaba et al., *Journal of Experimental Medicine* 173:549-559, 1991; Macatonia et al., *Journal of Immunology* 154:5071-5079, 1995; Porgador et al., *Journal of Experimental Medicine* 182:255-260, 1995; 15 Nair et al., *Journal of Virology* 67:4062-4069, 1993; Huang et al., *Science* 264:961-965, 1994; Macatonia et al., *Journal of Experimental Medicine* 169:1255-1264, 1989; Bhardwaj et al., *Journal of Clinical Investigation* 94:797-807, 1994; and Inaba et al., *Journal of Experimental Medicine* 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, 20 proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., *Cytometry* 13:795-808, 1992; Gorczyca et al., *Leukemia* 7:659-670, 1993; Gorczyca et al., *Cancer Research* 53:1945-1951, 1993; Itoh et al., *Cell* 66:233-243, 1991; Zacharchuk, *Journal of Immunology* 145:4037-4045, 1990; Zamai et al., *Cytometry* 14:891-897, 1993; 25 Gorczyca et al., *International Journal of Oncology* 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., *Blood* 84:111-117, 1994; Fine et al., *Cellular Immunology* 155:111-122, 1994; Galy et al., *Blood* 85:2770-2778, 1995; Toki et al., *Proc. Nat. Acad Sci. USA* 88:7548-7551, 1991.

30

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell

lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and

Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., *Experimental Hematology* 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long
5 term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland, H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

10

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns,
15 incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as
20 well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal
25 disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue
30 destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in

circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-

β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., *Endocrinology* 91:562-572, 1972; Ling et al., *Nature* 321:779-782, 1986; Vale et al., *Nature* 321:776-779, 1986; Mason et al., *Nature* 318:659-663, 1985; Forage et al., *Proc. Natl. Acad. Sci. USA* 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion

- include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

- A protein of the invention may also exhibit hemostatic or thrombolytic activity.
- 10 As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting
- 15 therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

- Assay for hemostatic and thrombolytic activity include, without limitation, those
- 20 described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

- 25 A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation,
- 30 cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without

limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

- 5 Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; 10 Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

- 15 Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or 20 suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin 25 lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

30 Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human

diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved
5 extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only
10 itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with
15 polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion
20 suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

25 Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and
30 reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from

forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

- 5 Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present
- 10 invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

- Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995;
- 15 Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

- In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities.
- 20 A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating
- 25 or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

- A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious
- 30 agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms;

effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

15 ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects

of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active

ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein

of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1mg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such

antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in
5 R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where
10 abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage,
15 tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or
20 tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the
25 composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

30 The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and

polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- α and TGF- β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
- (ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
- (iii) NUMBER OF SEQUENCES: 33
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Genetics Institute, Inc.
 - (B) STREET: 87 CambridgePark Drive
 - (C) CITY: Cambridge
 - (D) STATE: MA
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 02140
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Sprunger, Suzanne A.
 - (B) REGISTRATION NUMBER: 41,323
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (617) 498-8284
 - (B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1433 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCCCGTGGTT ACACAGCTAA TAGGTGGTGG AGATGGAGAC AGAATTCAAA CCCAGGCATT 60
CTTGATCTAC AGTATACACT CTTACCCACC ATCCTACACA GCCTTCTTA TTCATAAAAT 120
ATTTTCTACA GTGCAAGAAA ATTTTGATAG CTTGCTTATT TATTCAAGAT TTAGACTATA 180
TAGATTA ACT AGACTATCAA GATTTTAAAT TCTTGTTT TTTGTTTTTY YCCCCCTCTG 240
TGGCATAACT ATCTCTTAGT GATTTGAAGT TCTGATAGGC ATTTATTTAT GTTTTTGATT 300
AATTAAAAAA AGGGAAAAAA ATGGAACATA ATTATTGAAG CTATCGTCTA GGTA AAAAACC 360
TTTCTAAATG TAAGGTTTCAT TTAGATTGAT GACCTGTAGA GTGTAACAGT ATTGCCATAG 420
GCATACAGCT TTTTAATCAC ATATCATACA TAAACAAATT AGTAATACAG GTGGGTAGAT 480
ACAGACCCTA ACTTTGAGCT CTAAGATGAA ATTTGTTTAT AAATCCCTAG TTTCCATTCA 540
GTTTTTTCAA TATTTATCAA ACACCTACTG TGCCAGGCAT TGTTTAGGCA CAGGGGATAC 600
AGCAGGAGAA CAAAATGAAC AAAATTTTTT GCCTTCACAG AGCTAATTTT TTGTATTTTT 660
TTGTAGAGAT GGGGTTTTGC CATGTTTGCC AGTCTGGTCT CAACCTCCTA AGCTCAAGCA 720
GCCCACCCTC CTTGGCTTCC CAAAGTGCTG AGATTACAGG CATGAGCCAC CGCACTCTTC 780
TTAGCTATTT TTCATAGAAA CTTTATGTAT AAAAATAGAA GGGTAATGAC ACACCACCTT 840
TCTACTGATC TCCCCACTTC AGTAGTTATC ACATAACAGT CTTTTTTCAC CTATCTCCTT 900
CACTTTACCT CCTCTCCCTT AGTACTTTGA AGTAAATCTC AATGCAAGCT GGTATGTTTT 960
TCAAAATGAA ACATATAAAC ATGGACTAGA AAAAAATCTC TTCATACAGG ATTTGGTTTT 1020
GCAGAGAATT TACAAAGTGC GGTTAATGTA TGCCAATGGT TTCTCAGTTT GGATATCGAG 1080
ATCCTTAGAT GGACCATGAA GCTGGTAATA ATTTTATAGC TAACTTTTGT TAAGTGCTTA 1140
CTATATGCCA GGCAGTGTTC TAAGCATTTT ACGTGATTC ATTCATTCAG TTCTCACAAC 1200
TCTTTTAATT AGGTATTATT ATGATCTCCA TCTCAAAACA AAACAAAACA AAAAAATTAG 1260
CCTGGCATGG TGGCAGGCGC CTGTAATCCC AGTTACTTGA GAGGCTAAGG CAGGAGAATC 1320
GCTTGAATCT GGGAGGCAGA GGTTCAGTG AGCCGAGATT GCACTACTGC ACTCCAGCCT 1380
GGGTGACAGA ATGAGACTCT GTCTCAAAAA AAAAAAAAAA AAAAAAAAAA AAA 1433

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 46 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Lys Phe Val Tyr Lys Ser Leu Val Ser Ile Gln Phe Phe Gln Tyr
 1             5             10             15

Leu Ser Asn Thr Tyr Cys Ala Arg His Cys Leu Gly Thr Gly Asp Thr
      20             25             30

Ala Gly Glu Gln Asn Glu Gln Asn Phe Leu Pro Ser Gln Ser
      35             40             45

```

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1401 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

TCGGGACAGA TTTAAGTGCA GCGTGGATTT TTTTTTCTC ACTTTGCCTT GTGTTTTCCA      60
CCCTGAAAGA ATGTTGTGGC TGCTCTTTTT TCTGGTGA CTGCAATTCATG CTGAACTCTG      120
TCAACCAGGT GCAGAAAATG CTTTAAAGT GAGACTTAGT ATCAGAACAG CTCTGGGAGA      180
TAAAGCATAT GCCTGGGATA CCAATGAAGA ATACCTCTTC AAAGCGATGG TAGCTTTCTC      240
CATGAGAAAA GTTCCCAACA GAGAAGCAAC AGAAATTTCC CATGTCCTAC TTTGCAATGT      300
AACCCAGAGG GTATCATTCT GGTGTGTGGT TACAGACCCT TCAAAAATC ACACCCTTCC      360
TGCTGTTGAG GTGCAATCAG CCATAAGAAAT GAACAAGAAC CGGATCAACA ATGCCTTCTT      420
TGTAATGAC CAACTCTGG AATTTTAAA AATCCCTTCC AACTTGCAC CACCCATGGA      480

```


CCCATCTGTG CCCATCTGGA TTATTATATT TGGTGTGATA TTTTGCATCA TCATAGTTGC 540
 AATTGCACTA CTGATTTTAT CAGGGATCTG GCAACGTAGA AGAAAGAACA AAGAACCATC 600
 TGAAGTGGAT GACGCTGAAG ATAAGTGTGA AAACATGATC ACAATTGAAA ATGGCATCCC 660
 CTCTGATCCC CTGGACATGA AGGGAGGGCA TATTAATGAT GCCTTCATGA CAGAGGATGA 720
 GAGGCTCACC CCTCTCTGAA GGGCTGTTGT TCTGCTTCCT CAAGAAATTA AACATTTGTT 780
 TCTGTGTGAC TGCTGAGCAT CCTGAAATAC CAAGAGCAGA TCATATATTT TGTTCACCA 840
 TTCTTCTTTT GTAATAAATT TTGAATGTGC TTGAAAGTGA AAAGCAATCA ATTATACCCA 900
 CCAACACCAC TGAAATCATA AGCTATTAC GACTCAAAAT ATTCTAAAAT ATTTTCTGA 960
 CAGTATAGTG TATAAATGTG GTCATGTGGT ATTTGTAGTT ATTGATTAA GCATTTTGTAG 1020
 AAATAAGATC AGGCATATGT ATATATTTTC ACACTTCAA GACCTAAGGA AAAATAAATT 1080
 TTCCAGTGGG GAATACATAT AATATGGTGT AGAAATCATT GAAAATGGAT CCTTTTGTGAC 1140
 GATCACTTAT ATCACTCTGT ATATGACTAA GTAAACAAAA GTGAGAAGTA ATTATTGTAA 1200
 ATGGATGGAT AAAAATGGAA TTAATCATAT ACAGGGTGGG ATTTTATCCT GTTATCACAC 1260
 CAACAGTTGA TTATATATTT TCTGAATATC AGCCCCTAAT AGGACAATTC TATTGTTGA 1320
 CCATTCTAC AATTTGTAAA AGTCCAATCT GTGCTAACTT AATAAAGTAA TAATCATCTC 1380
 TTTTAAAAAA AAAAAAAAAA A 1401

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 222 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Leu Trp Leu Leu Phe Phe Leu Val Thr Ala Ile His Ala Glu Leu
 1 5 10 15
 Cys Gln Pro Gly Ala Glu Asn Ala Phe Lys Val Arg Leu Ser Ile Arg
 20 25 30
 Thr Ala Leu Gly Asp Lys Ala Tyr Ala Trp Asp Thr Asn Glu Glu Tyr
 35 40 45

Leu Phe Lys Ala Met Val Ala Phe Ser Met Arg Lys Val Pro Asn Arg
 50 55 60
 Glu Ala Thr Glu Ile Ser His Val Leu Leu Cys Asn Val Thr Gln Arg
 65 70 75 80
 Val Ser Phe Trp Phe Val Val Thr Asp Pro Ser Lys Asn His Thr Leu
 85 90 95
 Pro Ala Val Glu Val Gln Ser Ala Ile Arg Met Asn Lys Asn Arg Ile
 100 105 110
 Asn Asn Ala Phe Phe Val Asn Asp Gln Thr Leu Glu Phe Leu Lys Ile
 115 120 125
 Pro Ser Thr Leu Ala Pro Pro Met Asp Pro Ser Val Pro Ile Trp Ile
 130 135 140
 Ile Ile Phe Gly Val Ile Phe Cys Ile Ile Ile Val Ala Ile Ala Leu
 145 150 155 160
 Leu Ile Leu Ser Gly Ile Trp Gln Arg Arg Arg Lys Asn Lys Glu Pro
 165 170 175
 Ser Glu Val Asp Asp Ala Glu Asp Lys Cys Glu Asn Met Ile Thr Ile
 180 185 190
 Glu Asn Gly Ile Pro Ser Asp Pro Leu Asp Met Lys Gly Gly His Ile
 195 200 205
 Asn Asp Ala Phe Met Thr Glu Asp Glu Arg Leu Thr Pro Leu
 210 215 220

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 441 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCGGCCGCAG GTCTAGAATT CAATCGGCCA CAAGCTACTC TTTGGAGCCC ATCTATGGTT 60
 TGTGGTATGA CCACTCCTCC AACTTCTCCT GGAAATGTCC CACCTGATCT GTCACACCCT 120
 TACAGTAAAG TCTTTGGTAC AACTGCAGGT GGAAAAGGAA CTCCTCTGGG AACCCCAGCA 180
 ACCTCTCCTC CTCCAGCCCC ACTCTGTCAT TCGGATGACT ACGTGACAT TCACTCCCC 240

CAGGCCACAG TCACACCCCC CAGGAAGGAA GAGAGAATGG ATTCTGCAAG ACCATGTCTA 300
 CACAGACAAC ACCATCTTCT GAATGACAGA GGATCAGAAG AGCCACCTGG CAGCAAAGGT 360
 TCTGTCACTC TAAGTGATCT TCCAGGGTTT TTAGGTGATC TGGCCTCTGA AGAAGATAGT 420
 ATTGAAAAAA AAAAAAAAAA A 441

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Val	Cys	Gly	Met	Thr	Thr	Pro	Pro	Thr	Ser	Pro	Gly	Asn	Val	Pro	1	5	10	15
Pro	Asp	Leu	Ser	His	Pro	Tyr	Ser	Lys	Val	Phe	Gly	Thr	Thr	Ala	Gly	20	25	30	
Gly	Lys	Gly	Thr	Pro	Leu	Gly	Thr	Pro	Ala	Thr	Ser	Pro	Pro	Pro	Ala	35	40	45	
Pro	Leu	Cys	His	Ser	Asp	Asp	Tyr	Val	His	Ile	Ser	Leu	Pro	Gln	Ala	50	55	60	
Thr	Val	Thr	Pro	Pro	Arg	Lys	Glu	Glu	Arg	Met	Asp	Ser	Ala	Arg	Pro	65	70	75	80
Cys	Leu	His	Arg	Gln	His	His	Leu	Leu	Asn	Asp	Arg	Gly	Ser	Glu	Glu	85	90	95	
Pro	Pro	Gly	Ser	Lys	Gly	Ser	Val	Thr	Leu	Ser	Asp	Leu	Pro	Gly	Phe	100	105	110	
Leu	Gly	Asp	Leu	Ala	Ser	Glu	Glu	Asp	Ser	Ile	115	120							

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2353 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAGAAGGCGA TGTCACTATT GGAGAAGATG CACCAAATCT TTCTTTTAGC ACCAGTGTGG 60
GAAATGAGGA CGCCAGGACA GCCTGGCCCG AATTACAACA GAGCCATGCT GTTAATCAGC 120
TCAAAGATTT GTTGCGCCAA CAAGCAGATA AGGAAAAGTGA AGTATCTCCG TCAAGAAGAA 180
GAAAAATGTC CCCCTTGAGG TCATTAGAAC ATGAGGAAAC CAATATGCCT ACTATGCACG 240
ACCTTGTTCA TACTATTAAT GACCAGTCTC AATATATTCA TCATTTAGAG GCAGAAGTTA 300
AGTTCTGCAA GGAGGAACTC TCTGGAATGA AAAATAAAAT ACAAGTAGTT GTGCTTGAAA 360
ACGAAGGGCT CCAGCAACAG CTAAAATCTC AAAGACAAGA GGAGACACTG AGGGAACAAA 420
CACTTCTGGA TGCATCCGGA AACATGCACA ATTCTTGAT TACAACAGGT GAAGATTCTG 480
GGGTGGGCGA AACCTCCAAA AGACCATTTT CCCATGACAA TGCAGATTTT GGCAAAGCTG 540
CATCTGCTGG TGAGCAGCTA GAACTGGAGA AGCTAAAACT TACTTATGAG GAAAAGTGTG 600
AAATTGAGGA ATCCCAATTG AAGTTTTTGA GGAACGACTT AGCTGAATAT CAGAGAACTT 660
GTGAAGATCT TAAAGAGCAA CTAAAGCATA AAGAATTTCT TCTGGCTGCT AATACTTGTA 720
ACCGTGTTGG TGGTCTTTGT TTGAAATGTG CTCAGCATGA AGCTGTTCTT TCCCAAACCC 780
ATACTAATGT TCATATGCAG ACCATCGAAA GACTGGTTAA AGAAAGAGAT GACTTGATGT 840
CTGCACTAGT TTCCGTAAGG AGCAGCTTGG CAGATACGCA GCAAAGAGAA GCAAGTGCTT 900
ATGAACAGGT GAAACAAGTT TTGCAAATAT CTGAGGAAGC CAATTTTGAA AAAACCAAGG 960
CTTTAATCCA GTGTGACCAG TTGAGGAAGG AGCTGGAGAG GCAGGCGGAG CGACTTGAAA 1020
AAGAACTTGC ATCTCAGCAA GAGAAAAGGG CCATTGAGAA AGACATGATG AAAAAGGAAA 1080
TAACGAAAGA AAGGGAGTAC ATGGGATCAA AGATGTTGAT CTTGTCTCAG AATATTGCCC 1140
AACTGGAGGC CCAGGTGGAA AAGGTTACAA AGGAAAAGAT TTCAGCTATT AATCAACTGG 1200
AGGAAATTCA AAGCCAGCTG GCTTCTCGGG AAATGGATGT CACAAAGGTG TGTGGAGAAA 1260
TGCGCTATCA GCTGAATAAA ACCAACATGG AGAAGGATGA GGCAGAAAAG GAGCACAGAG 1320
AGTTCAGAGC AAAAATAAC AGGGATCTTG AAATTAAAGA TCAGGAAATA GAGAAATTGA 1380
GAATAGAACT GGATGAAAGC AAACAACACT TGGAACAGGA GCAGCAGAAG GCAGCCCTGG 1440

CCAGAGAGGA GTGCCTGAGA CTAACAGAAC TGCTGGGCGA ATCTGAGCAC CAACTGCACC 1500
 TCACCAGACA GGAAAAAGAT AGCATTCAGC AGAGCTTTAG CAAGGAAGCA AAGGCCCAAG 1560
 CCCTTCAGGC CCAGCAAAGA GAGCAGGAGC TGACACAGAA GATACAGCAA ATGGAAGCCC 1620
 AGCATGACAA AACTGAAAAT GAACAGTATT TGTGCTGAC CTCCCAGAAT ACATTTTGA 1680
 CAAAGTTAAA GGAAGAATGC TGTACATTAG CCAAGAACT GGAACAAATC TCTCAAAAAA 1740
 CCAGATCTGA AATAGCTCAA CTCAGTCAAG AAAAAAGGTA TACATATGAT AAATTGGGAA 1800
 AGTTACAGAG AAGAAATGAA GAATTGGAGG AACAGTGTGT CCAGCATGGG AGAGTACATG 1860
 AGACGATGAA GCAAAGGCTA AGGCAGCTGG ATAAGCACAG CCAGGCCACA GCCCAGCAGC 1920
 TGGTGCAGCT CCTCAGCAAG CAGAACCAGC TTCTCCTGGA GAGGCAGAGC CTGTCGGAAG 1980
 AGGTGGACCG GCTGCGGACC CAGTTACCCA GCATGCCACA ATCTGATTGC TGACCTGGAT 2040
 GGAACAGAGT GAAATAAATG ATTTACAAAG AGATATTTAC ATTCATCTGG TTTAGACTTA 2100
 ATATGCCACA ACGCACACAG ACCTTCCCAG GGTGACACCG CCTCAGCCTG CAGTGGGGCT 2160
 GGTCTCATC AACGCGGGCG CTGTCCCCGC ACGCAGTCGG GCTGGAGCTG GAGTCTGACT 2220
 CTAGCTGAGC AGAGCTCCTG GTGTATGTTT TCAGAAATGG CTTGAAGTTA TGTGTTTAAA 2280
 TCTGCTCATT CGTATGCTAG GTTATACATA TGATTTTCAA TAAATGAACT TTTTAAAGAA 2340
 AAAAAAAAAA AAA 2353

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 615 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Ser	Pro	Leu	Arg	Ser	Leu	Glu	His	Glu	Glu	Thr	Asn	Met	Pro	Thr
1				5				10					15		
Met	His	Asp	Leu	Val	His	Thr	Ile	Asn	Asp	Gln	Ser	Gln	Tyr	Ile	His
			20					25					30		
His	Leu	Glu	Ala	Glu	Val	Lys	Phe	Cys	Lys	Glu	Glu	Leu	Ser	Gly	Met
			35				40					45			

Lys Asn Lys Ile Gln Val Val Val Leu Glu Asn Glu Gly Leu Gln Gln
 50 55 60
 Gln Leu Lys Ser Gln Arg Gln Glu Glu Thr Leu Arg Glu Gln Thr Leu
 65 70 75 80
 Leu Asp Ala Ser Gly Asn Met His Asn Ser Trp Ile Thr Thr Gly Glu
 85 90 95
 Asp Ser Gly Val Gly Glu Thr Ser Lys Arg Pro Phe Ser His Asp Asn
 100 105 110
 Ala Asp Phe Gly Lys Ala Ala Ser Ala Gly Glu Gln Leu Glu Leu Glu
 115 120 125
 Lys Leu Lys Leu Thr Tyr Glu Glu Lys Cys Glu Ile Glu Glu Ser Gln
 130 135 140
 Leu Lys Phe Leu Arg Asn Asp Leu Ala Glu Tyr Gln Arg Thr Cys Glu
 145 150 155 160
 Asp Leu Lys Glu Gln Leu Lys His Lys Glu Phe Leu Leu Ala Ala Asn
 165 170 175
 Thr Cys Asn Arg Val Gly Gly Leu Cys Leu Lys Cys Ala Gln His Glu
 180 185 190
 Ala Val Leu Ser Gln Thr His Thr Asn Val His Met Gln Thr Ile Glu
 195 200 205
 Arg Leu Val Lys Glu Arg Asp Asp Leu Met Ser Ala Leu Val Ser Val
 210 215 220
 Arg Ser Ser Leu Ala Asp Thr Gln Gln Arg Glu Ala Ser Ala Tyr Glu
 225 230 235 240
 Gln Val Lys Gln Val Leu Gln Ile Ser Glu Glu Ala Asn Phe Glu Lys
 245 250 255
 Thr Lys Ala Leu Ile Gln Cys Asp Gln Leu Arg Lys Glu Leu Glu Arg
 260 265 270
 Gln Ala Glu Arg Leu Glu Lys Glu Leu Ala Ser Gln Gln Glu Lys Arg
 275 280 285
 Ala Ile Glu Lys Asp Met Met Lys Lys Glu Ile Thr Lys Glu Arg Glu
 290 295 300
 Tyr Met Gly Ser Lys Met Leu Ile Leu Ser Gln Asn Ile Ala Gln Leu
 305 310 315 320
 Glu Ala Gln Val Glu Lys Val Thr Lys Glu Lys Ile Ser Ala Ile Asn
 325 330 335
 Gln Leu Glu Glu Ile Gln Ser Gln Leu Ala Ser Arg Glu Met Asp Val

340

345

350

Thr Lys Val Cys Gly Glu Met Arg Tyr Gln Leu Asn Lys Thr Asn Met
 355 360 365
 Glu Lys Asp Glu Ala Glu Lys Glu His Arg Glu Phe Arg Ala Lys Thr
 370 375 380
 Asn Arg Asp Leu Glu Ile Lys Asp Gln Glu Ile Glu Lys Leu Arg Ile
 385 390 395 400
 Glu Leu Asp Glu Ser Lys Gln His Leu Glu Gln Glu Gln Lys Ala
 405 410 415
 Ala Leu Ala Arg Glu Glu Cys Leu Arg Leu Thr Glu Leu Leu Gly Glu
 420 425 430
 Ser Glu His Gln Leu His Leu Thr Arg Gln Glu Lys Asp Ser Ile Gln
 435 440 445
 Gln Ser Phe Ser Lys Glu Ala Lys Ala Gln Ala Leu Gln Ala Gln Gln
 450 455 460
 Arg Glu Gln Glu Leu Thr Gln Lys Ile Gln Gln Met Glu Ala Gln His
 465 470 475 480
 Asp Lys Thr Glu Asn Glu Gln Tyr Leu Leu Leu Thr Ser Gln Asn Thr
 485 490 495
 Phe Leu Thr Lys Leu Lys Glu Glu Cys Cys Thr Leu Ala Lys Lys Leu
 500 505 510
 Glu Gln Ile Ser Gln Lys Thr Arg Ser Glu Ile Ala Gln Leu Ser Gln
 515 520 525
 Glu Lys Arg Tyr Thr Tyr Asp Lys Leu Gly Lys Leu Gln Arg Arg Asn
 530 535 540
 Glu Glu Leu Glu Glu Gln Cys Val Gln His Gly Arg Val His Glu Thr
 545 550 555 560
 Met Lys Gln Arg Leu Arg Gln Leu Asp Lys His Ser Gln Ala Thr Ala
 565 570 575
 Gln Gln Leu Val Gln Leu Leu Ser Lys Gln Asn Gln Leu Leu Leu Glu
 580 585 590
 Arg Gln Ser Leu Ser Glu Glu Val Asp Arg Leu Arg Thr Gln Leu Pro
 595 600 605
 Ser Met Pro Gln Ser Asp Cys
 610 615

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 313 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
GCGACCTCTT CTGCGGCCCG CCTGGGCAGG TGTCTTCCTC GAGAGGCAGG CAGGGGATCC      60
CGGACACTAG CTTTATCGTC ATCTGGGAAA TTGTTAAAAA TGCAAATTCG CAAGTTTGAG      120
AGCCATGGTT CCAAGAAACT GCATAAGCAT ACGAAATAAG TTGCAGCCTC CCGACTTATA      180
CCCTGGTACT TCTAGTCTAA AACAGGATTT GACTCTACTA ATCCAGCCTT ATACAGGATG      240
CTGTGTTCTT TGCTCCTTTG TGAATGTCTG TTGCTGGTAG CTGGTTATGC TCATGATGAT      300
GACTGGATTG ACC                                     313
```

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 677 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
CCTTGATGA TGCATTAAGT GATATTTTAA TTAATTTTAA GTTTCATGAT TTTGAAACAT      60
GGAAGTGGCG ATTCGAAGAT TCCTTTGGAG TGGATCCATA TAATGTGTTA ATGGTAATTC      120
TTTGTCTGCT CTGCATCGTG GTTTTAGTGG CTA CTGAGCT GTGGACATAT GTATGTTGGT      180
ACACTCAGTT GAGACGTGTT TTAATCATCA GCTTTCTGTT CAGTTTGGGA TGGAATTGGA      240
TGTATTTATA TAAGCTAGCT TTTGCACAGC ATCAGGCTGA AGTCGCCAAG ATGGAGCCAT      300
TAAACAATGT GTGTGCCAAA AAGATGGACT GGACTGGAAG TATCTGGGAA TGGTTTAGAA      360
GTTTCATGGAC CTATAAGGAT GACCCATGCC AAAAATACTA TGAGCTCTTA CTAGTCAACC      420
CTATTTGGTT GGTCCCACCA ACAAAGGCAC TTGCAGTTAC ATTCACCACA TTTGTAACGG      480
```


AGCCATTGAA GCATATTGGA AAAGGAACTG GGGAATTTAT TAAAGCACTC ATGAAGGAAA 540
 TTCCAGCGCT GCTTCATCTT CCAGTGCTGA TAATTATGGC ATTAGCCATC CTGAGTTTCT 600
 GCTATGGTGC TGGAAAATCA GTTCATGTGC TGAGACATAT AGGCGGTCCT GAGAGCGAAC 660
 CTCCCCAGGC ACTTCGG 677

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 189 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Val Ile Leu Cys Leu Leu Cys Ile Val Val Leu Val Ala Thr Glu
 1 5 10 15
 Leu Trp Thr Tyr Val Cys Trp Tyr Thr Gln Leu Arg Arg Val Leu Ile
 20 25 30
 Ile Ser Phe Leu Phe Ser Leu Gly Trp Asn Trp Met Tyr Leu Tyr Lys
 35 40 45
 Leu Ala Phe Ala Gln His Gln Ala Glu Val Ala Lys Met Glu Pro Leu
 50 55 60
 Asn Asn Val Cys Ala Lys Lys Met Asp Trp Thr Gly Ser Ile Trp Glu
 65 70 75 80
 Trp Phe Arg Ser Ser Trp Thr Tyr Lys Asp Asp Pro Cys Gln Lys Tyr
 85 90 95
 Tyr Glu Leu Leu Leu Val Asn Pro Ile Trp Leu Val Pro Pro Thr Lys
 100 105 110
 Ala Leu Ala Val Thr Phe Thr Thr Phe Val Thr Glu Pro Leu Lys His
 115 120 125
 Ile Gly Lys Gly Thr Gly Glu Phe Ile Lys Ala Leu Met Lys Glu Ile
 130 135 140
 Pro Ala Leu Leu His Leu Pro Val Leu Ile Ile Met Ala Leu Ala Ile
 145 150 155 160
 Leu Ser Phe Cys Tyr Gly Ala Gly Lys Ser Val His Val Leu Arg His
 165 170 175

Ile Gly Gly Pro Glu Ser Glu Pro Pro Gln Ala Leu Arg
 180 185

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 470 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

AGACGGCAGG AGGAATTGAT TATAGACCTG ATGGTGGAGC AGGTGATGCC GATTTCATT      60
ATAGGGGCCA AATGGGCCCC ATTGAGCAAG GCCCTTATGC CAAAATGTAT GAGGGTAGAA      120
GAGAGATTTT GAGAGAGAGA GATGTTGACT TGAGATTTCA GGCTGGTCTC GAACTCCTGA      180
CCTCAAGTGA CCCGCCCTTG TCGGCCTCCC AAAGTGCTGG GATTACAGGC ATGAGCCATT      240
GTGCCCAGCC TATATAGTGT GAAGCTTTTA GGAAAATCAG AACAGGGTAG ACAGTTGTTA      300
AAAACAATGT TTAAATGGAA TAATGTTGAA TGTTTACAGG CTGTAAGAAT TATTGTATAC      360
ACAAAAATAAT ACACAAAGTT TGTACTTTGT GTACAAATAC AAATTTGTAC TTTGTGTACA      420
AATAATACAA AAAGTTTGTA TACACAAAAA AAAAAAAAAA AAAAAAAAAA      470

```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2702 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

CTGGAGTCCA CGCGGATTTT CGAAGCTGGG GCTGGCAAGA GGCCGCTGGA CACCACGCTC      60
CAGTCGTCAG CCCACTTCCT AGCTGAACAG CGCGAGGCGG CGGCAGCGAG CCGGGTCCCA      120
CCATGGCCGC GAATTATTCC AGTACCAGTA CCCGGAGAGA ACATGTCAAA GTTAAAACCA      180

```

GCTCCCAGCC AGGCTTCCTG GAACGGCTGA GCGAGACCTC GGGTGGGATG TTTGTGGGGC 240
TCATGGCCTT CCTGCTCTCC TTCTACCTAA TTTTCACCAA TGAGGGCCGC GCATTGAAGA 300
CGGCAACCTC ATTGGCTGAG GGGCTCTCGC TTGTGGTGTC TCCTGACAGC ATCCACAGTG 360
TGGCTCCGGA GAATGAAGGA AGGCTGGTGC ACATCATTGG CGCCTTACGG ACATCCAAGC 420
TTTTGTCTGA TCCAAACTAT GGGGTCCATC TTCCGGCTGT GAAACTGCGG AGGCACGTGG 480
AGATGTACCA ATGGGTAGAA ACTGAGGAGT CCAGGGAGTA CACCGAGGAT GGGCAGGTGA 540
AGAAGGAGAC GAGGTATTCC TACAACACTG AATGGAGGTC AGAAATCATC AACAGCAAAA 600
ACTTCGACCG AGAGATTGGC CACAAAAACC CCAGCTTCCT CTCTCCCACA GTGCCATGGC 660
AGTGGAGTCA TTCATGGCAA CAGCCCCCTT TGTCCAAAT TGGCAGGTTTT TCCTCTCGTC 720
AGGCCTCATC GACAAAGTCG ACAACTTCAA GTCCCTGAGC CTATCCAAGC TGGAGGACCC 780
TCATGTGGAC ATCATTCGCC GTGGAGACTT TTTCTACCAC AGCGAAAATC CCAAGTATCC 840
AGAGGTGGGA GACTTGCGTG TCTCCTTTTC CTATGCTGGA CTGAGCGGCG ATGACCCTGA 900
CCTGGGCCCA GCTCACGTGG TAACCTGGCT TCCCAGGGGC AGACACTAAG TCAGAGCCTC 960
ACGACTTTCC TGGACACAGA CACCTTGGTC AATGTCAGGA GCGCTTGAC CCCCTTTTCC 1020
CTGGGGAAAG GCACACTCTC GCACACTC TCAGCCAGGC ACGCTTCTGA GCAGTTTCAG 1080
AGCTCCCATG TCCCCACAGC CATCCATGGA CCCCACGTTA AGAAGGGCAG CTCAAAAGGG 1140
GTCTCATAGT CGCACCTTAT GACAGGTGTT CCAGTCACAC ACAGACCCTC TCCCCAAGCC 1200
CGTTTTGATC TGTCAATAAT TGGTCTTGCG TTCCTGGCCT ATGTGCAGTC CTGCCCCATC 1260
CCCTGCTCTG CGCACTGCCC AAGAGCTTTG AATGCCTGGA GCTTTGAATG GAGCAGCTCA 1320
GCCAGAGCTG CAGAGGTGGA TGCATCCCAG ATGGATGTAT AGAGAGAGAA GCCCAGGGT 1380
CTCTGTGCTC ACTTCCCCAG CCGGCACCCA GTCCCGGGAG GGTGGGCCAT GGCTCTCATG 1440
GGCGTGTCTC CCGCTGGTCA CCCCTCAGCT CTAACACCAG GTCTCTGAC CAGGTCACTG 1500
TGATTGCCCCG GCAGCGGGGT GACCAGCTAG TCCCATTCTC CACCAAGTCT GGGGATACCT 1560
TACTGCTCCT GCACCACGGG GACTTCTCAG CAGAGGAGGT GTTTCATAGA GAACTAAGGA 1620
GCAACTCCAT GAAGACCTGG GGCCTGCGGG CAGCTGGCTG GATGGCCATG TTCATGGGCC 1680
TCAACCTTAT GACACGGATC CTCTACACCT TGGTGGACTG GTTTCCTGTT TTCCGAGACC 1740
TGGTCAACAT TGGCCTGAAA GCCTTTGCCT TCTGTGTGGC CACCTCGCTG ACCCTGCTGA 1800
CCGTGGCGGC TGGCTGGCTC TTCTACCGAC CCCTGTGGGC CCTCCTCATT GCCGGCCTGG 1860

```

CCCTTGTGCC CATCCTTGTT GCTCGGACAC GGGTGCCAGC CAAAAAGTTG GAGTGAAAAG 1920
ACCCTGGCAC CCGCCCGACA CCTGCGTGAG CCCTAGGATC CAGGTCCTCT CTCACCTCTG 1980
ACCCAGCTCC ATGCCAGAGC AGGAGCCCCG GTCAATTTTG GACTCTGCAC CCCCTCTCCT 2040
CTTCAGGGGC CAGACTTGGC AGCATGTGCA CCAGTTGGT GTTCACCAGC TCATGTCTTC 2100
CCCACATCTC TTCTTGCCAG TAAGCAGCTT TGGTGGGCAG CAGCAGCTCA TGAATGGCAA 2160
GCTGACAGCT TCTCCTGCTG TTTCCTTCCT CTCTTGACT GAGTGGGTAC GGCCAGCCAC 2220
TCAGCCCATT GGCAGCTGAC AACGCAGACA CGCTCTACGG AGGCCTGCTG ATAAAGGGCT 2280
CAGCCTTGCC GTGTGCTGCT TCTCATCACT GCACACAAGT GCCATGCTTT GCCACCACCA 2340
CCAAGCACAT CTGTGATCCT GAAGGGCGGC CGTTAGTCAT TACTGCTGAG TCCTGGGTCA 2400
CCAGCAGACA CACTGGGCAT GGACCCCTCA AAGCAGGCAC ACCCAAACA CAAGTCTGTG 2460
GCTAGAACCT GATGTGGTGT TAAAAGAGA AGAAACACTG AAGATGTCCT GAGGAGAAAA 2520
GCTGGACATA TACTGGGCTT CACACTTATC TTATGGCTTG GCAGAATCTT TGTAGTGTGT 2580
GGGATCTCTG AAGGCCCTAT TTAAGTTTTT CTCGTTACT TTGCTGCTTC ATGTGTACTT 2640
TCCTACCCCA AGAGGAAGTT TTCTGAAATA AGATTTAAAA ACAAACAAA AAAAAAAAAA 2700
AA 2702

```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

Met Ala Ala Asn Tyr Ser Ser Thr Ser Thr Arg Arg Glu His Val Lys
1             5             10             15
Val Lys Thr Ser Ser Gln Pro Gly Phe Leu Glu Arg Leu Ser Glu Thr
20             25             30
Ser Gly Gly Met Phe Val Gly Leu Met Ala Phe Leu Leu Ser Phe Tyr
35             40             45
Leu Ile Phe Thr Asn Glu Gly Arg Ala Leu Lys Thr Ala Thr Ser Leu

```

50 55 60
 Ala Glu Gly Leu Ser Leu Val Val Ser Pro Asp Ser Ile His Ser Val
 65 70 75 80
 Ala Pro Glu Asn Glu Gly Arg Leu Val His Ile Ile Gly Ala Leu Arg
 85 90 95
 Thr Ser Lys Leu Leu Ser Asp Pro Asn Tyr Gly Val His Leu Pro Ala
 100 105 110
 Val Lys Leu Arg Arg His Val Glu Met Tyr Gln Trp Val Glu Thr Glu
 115 120 125
 Glu Ser Arg Glu Tyr Thr Glu Asp Gly Gln Val Lys Lys Glu Thr Arg
 130 135 140
 Tyr Ser Tyr Asn Thr Glu Trp Arg Ser Glu Ile Ile Asn Ser Lys Asn
 145 150 155 160
 Phe Asp Arg Glu Ile Gly His Lys Asn Pro Ser Phe Leu Ser Pro Thr
 165 170 175
 Val Pro Trp Gln Trp Ser His Ser Trp Gln Gln Pro Pro Leu Ser Lys
 180 185 190
 Leu Ala Gly Phe Ser Ser Arg Gln Ala Ser Ser Thr Lys Ser Thr Thr
 195 200 205
 Ser Ser Pro
 210

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3395 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATCTTCCTGC CCTTCACCTG CATTGGCTAC ACGGCCACCA ATCAGGACTT CATCCAGCGC 60
 CTGAGCACAC TGATCCGGCA GGCCATCGAG CGGCAGCTGC CTGCCTGGAT CGAGGCTGCC 120
 AACCAGCGGG AGGAGGGCCA GGGTGAACAG GGCGAGGAGG AGGATGAGGA GGAGGAAGAA 180
 GAGGAGGACG TGGCTGAGAA CCGCTACTTT GAAATGGGGC CCCCAGACGT GGAGGAGGAG 240

GAGGGAGGAG GCCAGGGGGA GGAAGAGGAG GAGGAAGAGG ARGATGAAGA RGCCGAGGAG 300
GAGCGCCTGG CTCTGGAATG GGCCCTGGGC GCGGACGAAG ACTTCCTGCT GGAGCACATC 360
CGCATCCTCA AGGTGCTGTG GTGCTTCCTG ATCCATGTGC AGGGCAGTAT CCGCCAGTTC 420
GCCGCTGCC TTGTGCTCAC CGACTTCGGC ATCGCAGTCT TCGAGATCCC GCACCAGGAG 480
TCTCGGGGCA GCAGCCAGCA CATCCTCTCC TCCCTGCGCT TTGTCTTTTG CTCCCGCAT 540
GGCGACCTCA CCGAGTTTGG CTTCTCATG CCGGAGCTGT GTCTGGTGCT CAAGGTACGG 600
CACAGTGAGA ACACGCTCTT CATTATCTCG GACGCCGCCA ACCTGCACGA GTTCCACGG 660
GACCTGCGCT CATGCTTTGC ACCCCAGCAC ATGGCCATGC TGTGTAGCCC CATCCTCTAC 720
GGCAGCCACA CCAGCCTGCA GGAGTTCCTG CGCCAGCTGC TCACCTTCTA CAAGGTGGCT 780
GGCGGCTGCC AGGAGCGCAG CCAGGGCTGC TTCCCCGTCT ACCTGGTCTA CAGTGACAAG 840
CGCATGGTGC AGACGGCCGC CGGGGACTAC TCAGGCAACA TCGAGTGGGC CAGCTGCACA 900
CTCTGTTCAG CCGTGCGGCG CTCCTGCTGC GCGCCCTCTG AGGCCGTCAA GTCCGCCGCC 960
ATCCCCTACT GGCTGTTGCT CACGCCCCAG CACCTCAACG TCATCAAGGC CGACTTCAAC 1020
CCCATGCCCC ACCGTGGCAC CCACAACGTG CGCAACCGCA ACAGCTTCAA GCTCAGCCGT 1080
GTGCCGCTCT CCACCGTGCT GCTGGACCCC ACACGCAGCT GTACCCAGCC TCGGGGCGCC 1140
TTTGCTGATG GCCACGTGCT AGAGCTGCTC GTGGGGTACC GCTTTGTCAC TGCCATCTTC 1200
GTGCTGCCCC ACGAGAAGTT CCACTTCCTG CGCGTCTACA ACCAGCTGCG GGCCTCGCTG 1260
CAGGACCTGA AGACTGTGGT CATCGCCAAG ACCCCCAGGA CGGGAGGCAG CCCCCAGGGC 1320
TCCTTTGCGG ATGGCCAGCC TGCCGAGCGC AGGGCCAGCA ATGACCAGCG TCCCCAGGAG 1380
GTCCCAGCAG AGGCTCTGGC CCCGGCCCCA GTGGAAGTCC CAGCTCCAGC CCCTGCAGCA 1440
GCCTCAGCCT CAGGCCCAGC GAAGACTCCG GCCCCAGCAG AGGCCTCAAC TTCAGCTTTG 1500
GTCCCAGAGG AGACGCCAGT GGAAGCTCCA GCCCCACCCC CAGCCGAGGC CCCTGCCCAG 1560
TACCCGAGTG AGCACCTCAT CCAGGCCACC TCGGAGGAGA ATCAGATCCC CTCGCACTTG 1620
CCTGCCTGCC CGTCGCTCCG GCACGTCGCC AGCCTGCGGG GCAGCGCCAT CATCGAGCTC 1680
TTCCACAGCA GCATTGCTGA GGTTGAAAC GAGGAGCTGA GGCACCTCAT GTGGTCCTCG 1740
GTGGTGTTCT ACCAGACCCC AGGGCTGGAG GTGACTGCCT GCGTGCTGCT CTCCACCAAG 1800
GCTGTGTACT TTGTGCTCCA CGACGGCCTC CGCCGCTACT TCTCAGAGCC ACTGCAGGAT 1860
TTCTGGMATC AGAAAAACAC SGACTACAAC AACAGCCCTT TCCACATCTC CCAGTGCTTC 1920

GTGCTAAAGC TTAGTGACCT GCAGTCAGTC AATGTGGGGC TTTTCGACCA GCATTTCCGG 1980
CTGACGGGTT CCACCCCGAT GCAGGTGGTM ACGTGCTTGA CGCGGGACAG CTACCTGACG 2040
CACTGCTTCC TCCAGCACCT CATGGTCGTG CTGTCCTCTC TGGAACGCAC GCCCTCGCCG 2100
GAGCCTGTTG ACAAGGACTT CTA CTCTCCGAG TTTGGGAACA AGACCACAGG GAAGATGGAG 2160
AACTACGAGC TGATCCACTC TAGTCGCGTC AAGTTTACCT ACCCCAGTGA GGAGGAGATT 2220
GGGGACCTGA CGTTCACTGT GGCCCAAAAG ATGGCTGAGC CAGAGAAGGC CCCAGCCCTC 2280
AGCATCCTGC TGTACGTGCA GGCCTTCCAG GTGGGCATGC CACCCCTGG GTGCTGCAGG 2340
GGCCCCCTGC GCCCCAAGAC ACTCCTGCTC ACCAGCTCCG AGATCTTCCT CCTGGATGAG 2400
GACTGTGTCC ACTACCCACT GCCCGAGTTT GCCAAAGAGC CGCCGCAGAG AGACAGGTAC 2460
CGGCTGGACG ATGGCCGCCG CGTCCGGGAC CTGGACCGAG TGCTCATGGG CTACCAGACC 2520
TACCCGACCC CCCTCACCTT CGTCTTCGAT GACGTGCAAG GTCATGACCT CATGGGCAGT 2580
GTCACCCTGG ACCACTTTGG GGAGGTGCCA GGTGGCCCGG CTAGAGCCAG CCAGGGCCGT 2640
GAAGTCCAGT GGCAGGTGTT TGTCCCCAGT GCTGAGAGCA GAGAGAAGCT CATCTCGCTG 2700
TTGGCTCGCC AGTGGGAGGC CCTGTGTGGC CGTGAGCTGC CTGTCGAGCT CACCGGCTAG 2760
CCCAGGCCAC AGCCAGCCTG TCGTGTCCAG CCTGACGCCT ACTGGGGCAG GGCAGCAGGC 2820
TTTTGTGTTC TCTAAAAATG TTTTATCCTC CCTTTGGTAC CTTAATTTGA CTGTCCTCGC 2880
AGAGAATGTG AACATGTGTG TGTGTTGTGT TAATTCTTTC TCATGTTGGG AGTGAGAATG 2940
CCGGGCCCCCT CAGGGCTGTC GGTGTGCTGT CAGCCTCCCA CAGGTGGTAC AGCCGTGCAC 3000
ACCAGTGTCTG TGTCTGCTGT TGTGGGACCG TTGTTAACAC GTGACACTGT GGGTCTGACT 3060
TTCTCTTCTA CACGTCCTTT CCTGAAGTGT CGAGTCCAGT CCTTTGTTGC TGTGCTGTT 3120
GCTGTTGCTG TTGCTGTTGG CATCTTGCTG CTAATCCTGA GGCTGGTAGC AGAATGCACA 3180
TTGGAAGCTC CCACCCATA TTGTTCTTCA AAGTGGAGGT CTCCCCTGAT CCAGACAAGT 3240
GGGAGAGCCC GTGGGGGCAG GGGACCTGGA GCTGCCAGCA CCAAGCGTGA TTCCTGCTGC 3300
CTGTATTCTC TATTCCAATA AAGCAGAGTT TGACACCGTC AAAAAAAAAA AAAAAAAAAA 3360
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA 3395

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 848 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met	Gly	Pro	Pro	Asp	Val	Glu	Glu	Glu	Glu	Gly	Gly	Gly	Gln	Gly	Glu	1	5	10	15
Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Asp	Glu	Glu	Ala	Glu	Glu	Glu	Arg	Leu	20	25	30
Ala	Leu	Glu	Trp	Ala	Leu	Gly	Ala	Asp	Glu	Asp	Phe	Leu	Leu	Glu	His	35	40	45	
Ile	Arg	Ile	Leu	Lys	Val	Leu	Trp	Cys	Phe	Leu	Ile	His	Val	Gln	Gly	50	55	60	
Ser	Ile	Arg	Gln	Phe	Ala	Ala	Cys	Leu	Val	Leu	Thr	Asp	Phe	Gly	Ile	65	70	75	80
Ala	Val	Phe	Glu	Ile	Pro	His	Gln	Glu	Ser	Arg	Gly	Ser	Ser	Gln	His	85	90	95	
Ile	Leu	Ser	Ser	Leu	Arg	Phe	Val	Phe	Cys	Phe	Pro	His	Gly	Asp	Leu	100	105	110	
Thr	Glu	Phe	Gly	Phe	Leu	Met	Pro	Glu	Leu	Cys	Leu	Val	Leu	Lys	Val	115	120	125	
Arg	His	Ser	Glu	Asn	Thr	Leu	Phe	Ile	Ile	Ser	Asp	Ala	Ala	Asn	Leu	130	135	140	
His	Glu	Phe	His	Ala	Asp	Leu	Arg	Ser	Cys	Phe	Ala	Pro	Gln	His	Met	145	150	155	160
Ala	Met	Leu	Cys	Ser	Pro	Ile	Leu	Tyr	Gly	Ser	His	Thr	Ser	Leu	Gln	165	170	175	
Glu	Phe	Leu	Arg	Gln	Leu	Leu	Thr	Phe	Tyr	Lys	Val	Ala	Gly	Gly	Cys	180	185	190	
Gln	Glu	Arg	Ser	Gln	Gly	Cys	Phe	Pro	Val	Tyr	Leu	Val	Tyr	Ser	Asp	195	200	205	
Lys	Arg	Met	Val	Gln	Thr	Ala	Ala	Gly	Asp	Tyr	Ser	Gly	Asn	Ile	Glu	210	215	220	
Trp	Ala	Ser	Cys	Thr	Leu	Cys	Ser	Ala	Val	Arg	Arg	Ser	Cys	Cys	Ala	225	230	235	240

Pro Ser Glu Ala Val Lys Ser Ala Ala Ile Pro Tyr Trp Leu Leu Leu
 245 250 255
 Thr Pro Gln His Leu Asn Val Ile Lys Ala Asp Phe Asn Pro Met Pro
 260 265 270
 Asn Arg Gly Thr His Asn Cys Arg Asn Arg Asn Ser Phe Lys Leu Ser
 275 280 285
 Arg Val Pro Leu Ser Thr Val Leu Leu Asp Pro Thr Arg Ser Cys Thr
 290 295 300
 Gln Pro Arg Gly Ala Phe Ala Asp Gly His Val Leu Glu Leu Leu Val
 305 310 315 320
 Gly Tyr Arg Phe Val Thr Ala Ile Phe Val Leu Pro His Glu Lys Phe
 325 330 335
 His Phe Leu Arg Val Tyr Asn Gln Leu Arg Ala Ser Leu Gln Asp Leu
 340 345 350
 Lys Thr Val Val Ile Ala Lys Thr Pro Gly Thr Gly Gly Ser Pro Gln
 355 360 365
 Gly Ser Phe Ala Asp Gly Gln Pro Ala Glu Arg Arg Ala Ser Asn Asp
 370 375 380
 Gln Arg Pro Gln Glu Val Pro Ala Glu Ala Leu Ala Pro Ala Pro Val
 385 390 395 400
 Glu Val Pro Ala Pro Ala Pro Ala Ala Ala Ser Ala Ser Gly Pro Ala
 405 410 415
 Lys Thr Pro Ala Pro Ala Glu Ala Ser Thr Ser Ala Leu Val Pro Glu
 420 425 430
 Glu Thr Pro Val Glu Ala Pro Ala Pro Pro Pro Ala Glu Ala Pro Ala
 435 440 445
 Gln Tyr Pro Ser Glu His Leu Ile Gln Ala Thr Ser Glu Glu Asn Gln
 450 455 460
 Ile Pro Ser His Leu Pro Ala Cys Pro Ser Leu Arg His Val Ala Ser
 465 470 475 480
 Leu Arg Gly Ser Ala Ile Ile Glu Leu Phe His Ser Ser Ile Ala Glu
 485 490 495
 Val Glu Asn Glu Glu Leu Arg His Leu Met Trp Ser Ser Val Val Phe
 500 505 510
 Tyr Gln Thr Pro Gly Leu Glu Val Thr Ala Cys Val Leu Leu Ser Thr
 515 520 525
 Lys Ala Val Tyr Phe Val Leu His Asp Gly Leu Arg Arg Tyr Phe Ser

530 535 540
 Glu Pro Leu Gln Asp Phe Trp Xaa Gln Lys Asn Thr Asp Tyr Asn Asn
 545 550 555 560
 Ser Pro Phe His Ile Ser Gln Cys Phe Val Leu Lys Leu Ser Asp Leu
 565 570 575
 Gln Ser Val Asn Val Gly Leu Phe Asp Gln His Phe Arg Leu Thr Gly
 580 585 590
 Ser Thr Pro Met Gln Val Val Thr Cys Leu Thr Arg Asp Ser Tyr Leu
 595 600 605
 Thr His Cys Phe Leu Gln His Leu Met Val Val Leu Ser Ser Leu Glu
 610 615 620
 Arg Thr Pro Ser Pro Glu Pro Val Asp Lys Asp Phe Tyr Ser Glu Phe
 625 630 635 640
 Gly Asn Lys Thr Thr Gly Lys Met Glu Asn Tyr Glu Leu Ile His Ser
 645 650 655
 Ser Arg Val Lys Phe Thr Tyr Pro Ser Glu Glu Glu Ile Gly Asp Leu
 660 665 670
 Thr Phe Thr Val Ala Gln Lys Met Ala Glu Pro Glu Lys Ala Pro Ala
 675 680 685
 Leu Ser Ile Leu Leu Tyr Val Gln Ala Phe Gln Val Gly Met Pro Pro
 690 695 700
 Pro Gly Cys Cys Arg Gly Pro Leu Arg Pro Lys Thr Leu Leu Leu Thr
 705 710 715 720
 Ser Ser Glu Ile Phe Leu Leu Asp Glu Asp Cys Val His Tyr Pro Leu
 725 730 735
 Pro Glu Phe Ala Lys Glu Pro Pro Gln Arg Asp Arg Tyr Arg Leu Asp
 740 745 750
 Asp Gly Arg Arg Val Arg Asp Leu Asp Arg Val Leu Met Gly Tyr Gln
 755 760 765
 Thr Tyr Pro Gln Ala Leu Thr Leu Val Phe Asp Asp Val Gln Gly His
 770 775 780
 Asp Leu Met Gly Ser Val Thr Leu Asp His Phe Gly Glu Val Pro Gly
 785 790 795 800
 Gly Pro Ala Arg Ala Ser Gln Gly Arg Glu Val Gln Trp Gln Val Phe
 805 810 815
 Val Pro Ser Ala Glu Ser Arg Glu Lys Leu Ile Ser Leu Leu Ala Arg
 820 825 830

Gln Trp Glu Ala Leu Cys Gly Arg Glu Leu Pro Val Glu Leu Thr Gly
 835 840 845

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1147 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGAAGGAGTT CTGGAATTGG AAAACCATCA TTTTCAACC ATCACAGTAA ATATGGCTCA 60
 GGCAAGAATT ATCAATCAAT GCTAAAGCTA GGGGAAATT TCGCTTAGGA GCAGGATATT 120
 AGGGTATTAG TCTGGGCTTA AAGTATCTCC TCACAGATTG TTGTTAGTTT CTGGGGAAAG 180
 AATAGTAACC ATGCAATGGA AAAAAATGGA CAACCTCTTG ACTAGGTTAT CAAAATTAAC 240
 CTCACCAATA AAGGGTGGAT GTTCAACATG TGCCTTCAA TGTGACCCAC TGAGAAGGAA 300
 ACAACATCAC TGTAACAACA ACAACCAGAA ACGACAGGGG GTTTTGA CTG AATTCTTCAA 360
 AAATGTCAAT GTCATAGAAG ACAAAGAAAG GTTGTGGAAA TGTTCAGAT TAAATGATAG 420
 TAAAAACACC TGACAACTAA ACATAGTAAG TAATACTAGA CTGGATTCTG TACCAGAGGT 480
 AACATAAGTG CTCCAAAGGA CAATGTTAGG TCAACTGGCA AATTGGAATA TAGACAGTCA 540
 ATCAGATAAG AAGTATACTT TGATTAAGTA AAAAAATCC CTATTCTTGG AAAATACACA 600
 ATAAAGTATT TTGAGGTAAA GGGCCATAAT GTATGCAATC TACTCTCAA AAATTCAGAA 660
 ACATATATTT GTGTGCATTT GCATGTGCAA CAGTACACAC AAACATACAT AAAGAGAGCA 720
 ATTGATAAGG CAAATAAGGT AACATTTAAC AATAATCTGA TACACATAAA TAGAGAAAGA 780
 GCAATTGATA AAGTAAATGA GGTAAAATTT AACAATAATC TGAGCAAAAG GTATATGTGT 840
 TTTCTTTGAG ACAGTCTGAT TCTTGCAACT TATTCTGTAA GTTGGAAGTT ATTTCCAAAC 900
 ATGATTGAAA AAAACCCCG CACTTGGCAA CTTCTTCTCT TTTTCAGCCT AGAAATGTCT 960
 GTGTTAAGTG GTTTTTTATT TATTGTTGTT GTTGTGTTT ATTGTTGTTT TGTGCCAGG 1020
 CTCCAACCTCA CAAAATACGA GTTTAAAAAC TGCGTTGTTA TTTT TAGAGA TTTGTGATAA 1080

TACAACTTGT TATAAAATTT ATTCCTCAAT AAATATAATT TCTCTACTAT GCAAAAAAAA 1140
 AAAAAAA 1147

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Ile Glu Lys Lys Pro Arg Thr Trp Gln Leu Leu Leu Phe Phe Ser
 1 5 10 15
 Leu Glu Met Ser Val Leu Ser Gly Phe Leu Phe Ile Val Val Val Cys
 20 25 30
 Cys Tyr Cys Cys Phe Val Ala Arg Leu Gln Leu Thr Lys Tyr Glu Phe
 35 40 45
 Lys Asn Cys Val Val Ile Phe Arg Asp Leu
 50 55

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1013 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCTTTTAAAA AAATATCTGA AAAAAGCTTC ATATCTTTAC AAATCATAA AATAGCTGAT 60
 TGGGCCATGG AGGAGATGAG GCTGTTTAGA ACTGGTTTGT TTTCAAGTTT GTCAATTTTC 120
 CCTGTATGAG AACTTGGGTA AAGCACAAAG AACATACAG TGCTAGTAAC AGGTCCTCTG 180
 CGCCCTGGAA CTAAGTGTTT GGAGGAAGGA CTAAACCCCG GGGCAGGTGA GTATAAAATA 240
 ATTCCACTAA GATCACCTCC TCAGTCCCCA GAAGGCTGAT GGTGGATCCT CTGGCCATCT 300

```

CCTGTGGGGT CTTACTGCTC CTC TGCCATT TCTCTATGCC TGAAGACACG AAGATGATAT      360
CAAGGCAGAG CTACCATATC GCAGCCAGTC TCTAGGCTAC TGCTGTGCAG TGGCTCCAC      420
TTTCTAATGC TTTTTTGTTT TTGCTTTTC TAACAAAACA ATCTTTTTC AAAATGAATT      480
CCAACCCCTG CTAGTTCCTT CGCTGCCTCC ATACTGTTTT AGGCAGCACC GTTTATGTGA      540
CAGAGTCCGT GTTTCTCAAA TGCATGGTGT TCCTCAGGTG GAGAGTGGGC AGAAGTTTTT      600
GCAACACTTT TTTTTTAAGT TATTGGGTGC AAAATCCCAA ACCAGGATAT GTGTATGTCT      660
GTGTGTTTAT GTTTTTTATT TGACCCTCCC CTCTTCAAC CTACCCCTT TTATATCTAA      720
TGTAAGAAAA GCGAAATTGA ATCTGGAAG CAACTGTTG TATATAGTTG CGGTAACAAT      780
CATGAAGAGA GAGCCGGGCT GTCCCCTCAG TAATTCATTT TAAATAACAA ATTATTTAAA      840
AATAAAATTC ATGCCAGAGC CAGCTGAAGA GGCCTTCCTT CATCACCCT GAGGCCACCC      900
CCAATCTGGG CCCTCTCTCC ATCTGGCATG TCTCTCCCA GCAAGATTCA TCTGTTCAAT      960
GCCATTTGCG TTTCAATAAA GTTATCTCCT GTACTGTCAA AAAAAAAAAA AAA      1013

```

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 87 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Met His Gly Val Pro Gln Val Glu Ser Gly Gln Lys Phe Leu Gln His
1           5           10           15

Phe Phe Phe Lys Leu Leu Gly Ala Lys Ser Gln Thr Arg Ile Cys Val
20           25           30

Cys Leu Cys Val Tyr Val Phe Tyr Leu Thr Leu Pro Ser Phe Asn Leu
35           40           45

Pro Pro Phe Ile Ser Asn Val Glu Lys Ala Lys Leu Asn Leu Glu Ser
50           55           60

Lys Leu Leu Tyr Ile Val Ala Val Thr Ile Met Lys Arg Glu Pro Gly
65           70           75           80

Cys Pro Leu Ser Asn Ser Phe

```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1763 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TCGGGATAAA AAGCAAGAAA AGAAAGAGAA GACTGAGAAT AAGAAGATCT CTTTGAAAAT	60
AAAATAAGAC TGCTAAAAGT ATTTGGTATA CAGTCTGGAA AATAAAGTTG AGGGAATCTC	120
TCCAGATAAA GAGCAAAAAG AAATAGATAG AAAAATATAA AGAAAGAAAA AAGACATAGA	180
CAATCAATAT GTAATGTTAG GAGTTCCTGG AAGAGAGAAC AGAGACAGTG TAGGTGAAGA	240
AATAAAAAAGA AAAAGAATTG AAGAACAGAG CAAGCTAAGT CTCCAGATTG AGAGGGCCCA	300
ATACAATCTA CATCTAGACA CAATATTGTA AAATTTTGGA ATATTAAGGA TAGAAGGAAG	360
ATATTAAAGT GGCCAGGGAG AAAACAAATG AGGTCATCAC GATTAGCTCA ACACAAAAAT	420
GGATGAGAAA TAGACTGCTA ACAGATTTGT CATCAGCAAC ACTGAATGCC AGAAGTCAAT	480
GGATCAACAT CTTCAGAGCT TAAGGAAAAT TTTTGTACCT AGAATTTTCAT AGTAAGGCAG	540
ACTGTCAAGA AGAACATCAA AGTGAAGACA TTTTCTGTCA GGCAAATTTT CAGAAAGTCT	600
CCTTTGCACC CTTACTGAGG AAGTATCTTG AGGAAATTCT CCAGCAAAAT GAGGATGAAA	660
ACCAGGAAAG AAGAAGAAAT GGGATCCATA AAACAGTGGA CCTTACTTAG GATGTCTCAT	720
TCTAGAGTGA CAGCCAAAAG GGTATCTCAC CCTAGAGTGA CAGCTATCCA GCAGACTAAT	780
TTCAGATGAG AGCATACTGT CTCGGGCTTT CTGGGAAGAA TGTGCATTCA GTGCCATAGA	840
TAGTATCACT GAAGAGCTGG GATGCTTGAG AAGATTATTT AGTCAAGAAA AAAGAAAGAC	900
AAATCAACAA TATGTCAAAA AATTCAGGTC CAATTATAGA GCAAAATAAA ATGAGGCATG	960
ATTTTGAGTT ATTCATGAAG AATAAGAAGA GGCTTGATAG GTACATTTCC TTTTCTATGG	1020
CACAGGCATG ATGATATTGG GTGTGTAGGG AAGAAAATAT CCTAGCTTAT ACTAGGCTCC	1080
CAGTAAGAAG TATTTAAATA GCCAAAATAA TGTGGATATC ATTTATTAGT ATTCAATGTT	1140

CAGATCAGCC TATTAACAAA GTGTGAAAGG TTTCATTTTT TATTCAGAAC TGAAGTTGAA 1200
 AGTAATTAAT GCTGACAAAG GGAAAGAAAG CAGAAAGAGA TTGAGAATTA GAGGAAGAGA 1260
 AGTGGAATCA AAGGTAGAGA TACTTATATA TTCAAAGTGG GGATGAAAAG ATCTTCAGTT 1320
 AATGGAACAA GAACTAGAGG ATTAGTGTAT TGTTCAAAGC TATAAAATCA AACCAATAGA 1380
 TGTATTAAAA AGTGATGTAA CTATCAGACA TTTGGAGAGA GATGGACAAA GGAAAGTGGC 1440
 GATAGTGTAA GTTAAATCCT TATCTTTTGT AATGGGGAAT TATTAAAGAT GTTGTAAGT 1500
 CAGTAAGTCA AGAAATTATT GCTCAAACAT ATTATTTAAA GTTAGAAAGT TACCAGACGA 1560
 TCTAAAATAA ATATTGTAA AAGCATTACC TCTAGGGAAT GGGATTAGA TTTAAAAGG 1620
 GTGGGATGGG AAACGTGTT TTTCATTTTA AGTCCTTCTG TACTATTTAA TTTTTTACCT 1680
 TGTGCATGTA TTACTTTGAA AAAATTTTAA ATAAACCCAA ATAAAAATCT AAAAAAAAAA 1740
 AAAAAAAAAA AAAAAAAAAA AAA 1763

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met	Arg	Met	Lys	Thr	Arg	Lys	Glu	Glu	Glu	Met	Gly	Ser	Ile	Lys	Gln
1				5					10					15	
Trp	Thr	Leu	Leu	Arg	Met	Ser	His	Ser	Arg	Val	Thr	Ala	Lys	Arg	Val
			20					25					30		
Ser	His	Pro	Arg	Val	Thr	Ala	Ile	Gln	Gln	Thr	Asn	Phe	Arg		
		35					40					45			

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CNATCTTAGAG CTCAAAGTTA GGGTCTG

28

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CNCAGAGCTGT TCTGATACTA AGTCTCAC

29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ANACTATCTTC TTCAGAGGCC AGATCACC

29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CNAGAAGCCAG CTGGCTTTGA ATTCCTC

29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CNTTTTCCAAT ATGCTTCAAT GGCTCCGT

29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TNGGTAGAAGG AGAGCAGGAA GGCCATGA

29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GNCTTCTCTGG CTCAGCCATC TTTTGGGC

29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CNGTACACACA AACATACATA AAGAGAGC

29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ANACGGACTCT GTCACATAAA CGGTGCTG

29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GNTGAGATACC CTTTGGCTG TCACTCTA

29

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met	Gln	Trp	Lys	Lys	Met	Asp	Asn	Leu	Leu	Thr	Arg	Leu	Ser	Lys	Leu
1				5					10					15	
Thr	Ser	Pro	Ile	Lys	Gly	Gly	Cys	Ser	Thr	Cys	Ala	Phe	Lys	Cys	Asp
			20				25						30		
Pro	Leu	Arg	Arg	Lys	Gln	His	His	Cys	Asn	Asn	Asn	Asn	Gln	Lys	Arg
			35				40					45			
Gln	Gly	Val	Leu	Thr	Glu	Phe	Phe	Lys	Asn	Val	Asn	Val	Ile	Glu	Asp
	50					55				60					
Lys	Glu	Arg	Leu	Trp	Lys	Cys	Phe	Arg	Leu	Asn	Asp	Ser	Lys	Asn	Thr
65				70					75					80	

What is claimed is:

1. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 506 to nucleotide 643;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 471 to nucleotide 765;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AA35_2 deposited under accession number ATCC 98303;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AA35_2 deposited under accession number ATCC 98303;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AA35_2 deposited under accession number ATCC 98303;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AA35_2 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

2. A composition of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with a composition of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.
5. A process for producing a protein encoded by a composition of claim 2, which process comprises:
 - (a) growing a culture of the host cell of claim 3 in a suitable culture medium; and
 - (b) purifying said protein from the culture.
6. A protein produced according to the process of claim 5.
7. The protein of claim 6 comprising a mature protein.
8. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:
 - (a) the amino acid sequence of SEQ ID NO:2;
 - (b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32;
 - (c) fragments of the amino acid sequence of SEQ ID NO:2; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone AA35_2 deposited under accession number ATCC 98303;the protein being substantially free from other mammalian proteins.
9. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.
10. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32.
11. The composition of claim 8, further comprising a pharmaceutically acceptable carrier.
12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

14. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 71 to nucleotide 736;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 113 to nucleotide 736;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 343;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AM42_3 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AM42_3 deposited under accession number ATCC 98303;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AM42_3 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AM42_3 deposited under accession number ATCC 98303;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

15. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
 - (b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91;
 - (c) fragments of the amino acid sequence of SEQ ID NO:4; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone AM42_3 deposited under accession number ATCC 98303;
- the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

17. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 55 to nucleotide 423;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG137_7 deposited under accession number ATCC 98303;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG137_7 deposited under accession number ATCC 98303;
- (e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG137_7 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG137_7 deposited under accession number ATCC 98303;
- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;
- (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity;
- (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species' homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

18. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 62 to amino acid 123;

(c) fragments of the amino acid sequence of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone BG137_7 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

20. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 186 to nucleotide 2030;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 873 to nucleotide 2030;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 802 to nucleotide 1173;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CH699_1 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CH699_1 deposited under accession number ATCC 98303;

- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CH699_1 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CH699_1 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

21. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:8;
- (b) the amino acid sequence of SEQ ID NO:8 from amino acid 218 to amino acid 329;
- (c) fragments of the amino acid sequence of SEQ ID NO:8; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CH699_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

23. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 111 to nucleotide 677;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 156 to nucleotide 677;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO851_1 deposited under accession number ATCC 98303;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO851_1 deposited under accession number ATCC 98303;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO851_1 deposited under accession number ATCC 98303;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO851_1 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:11;

(b) the amino acid sequence of SEQ ID NO:11 from amino acid 120 to amino acid 189;

(c) fragments of the amino acid sequence of SEQ ID NO:11; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO851_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:10, SEQ ID NO:9 or SEQ ID NO:12 .

26. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 123 to nucleotide 755;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 279 to nucleotide 755;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 631;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CP111_1 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CP111_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CP111_1 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CP111_1 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

27. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- (b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 171;
- (c) fragments of the amino acid sequence of SEQ ID NO:14; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CP111_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.

29. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 214 to nucleotide 2760;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 406 to nucleotide 2760;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 2011 to nucleotide 2565;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CS278_1 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CS278_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CS278_1 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CS278_1 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

30. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:16;
 - (b) the amino acid sequence of SEQ ID NO:16 from amino acid 596 to amino acid 784;
 - (c) fragments of the amino acid sequence of SEQ ID NO:16; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone CS278_1 deposited under accession number ATCC 98303;
- the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

32. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 901 to nucleotide 1074;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 970 to nucleotide 1074;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 626 to nucleotide 1147;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DF968_3 deposited under accession number ATCC 98303;

- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DF968_3 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DF968_3 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DF968_3 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

33. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
- (b) fragments of the amino acid sequence of SEQ ID NO:18; and
- (c) the amino acid sequence encoded by the cDNA insert of clone DF968_3 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17.

35. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 560 to nucleotide 820;

- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DN1120_2 deposited under accession number ATCC 98303;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DN1120_2 deposited under accession number ATCC 98303;
- (e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DN1120_2 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DN1120_2 deposited under accession number ATCC 98303;
- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;
- (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity;
- (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;
- (j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and
- (k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

36. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:20;
- (b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 61;
- (c) fragments of the amino acid sequence of SEQ ID NO:20; and
- (d) the amino acid sequence encoded by the cDNA insert of clone DN1120_2 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:19.

38. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 649 to nucleotide 786;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 736 to nucleotide 786;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 525 to nucleotide 787;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DO589_1 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DO589_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DO589_1 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DO589_1 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:22;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:22 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

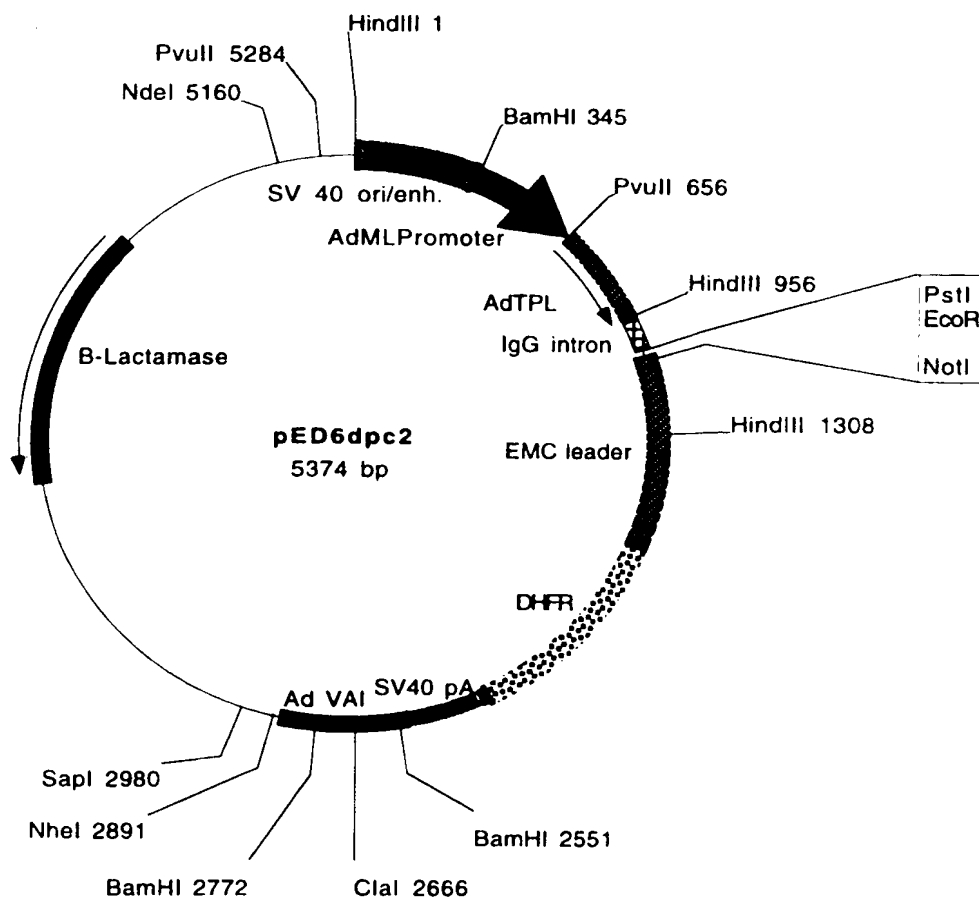
39. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:22;

- (b) fragments of the amino acid sequence of SEQ ID NO:22; and
- (c) the amino acid sequence encoded by the cDNA insert of clone DO589_1 deposited under accession number ATCC 98303; the protein being substantially free from other mammalian proteins.

40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:21.

FIGURE 1A

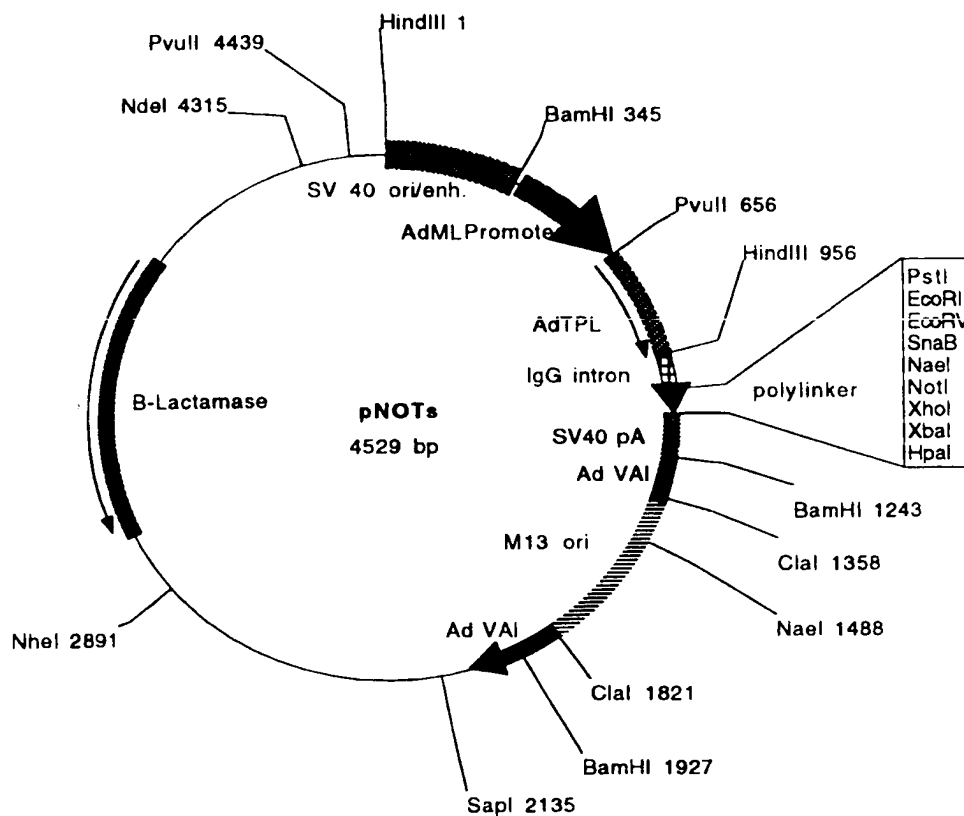


Plasmid name: pED6dpc2

Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs

Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al, 1989. Mol. Cell. Biol. 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the Clal site. SST cDNAs are cloned between EcoRI and NotI





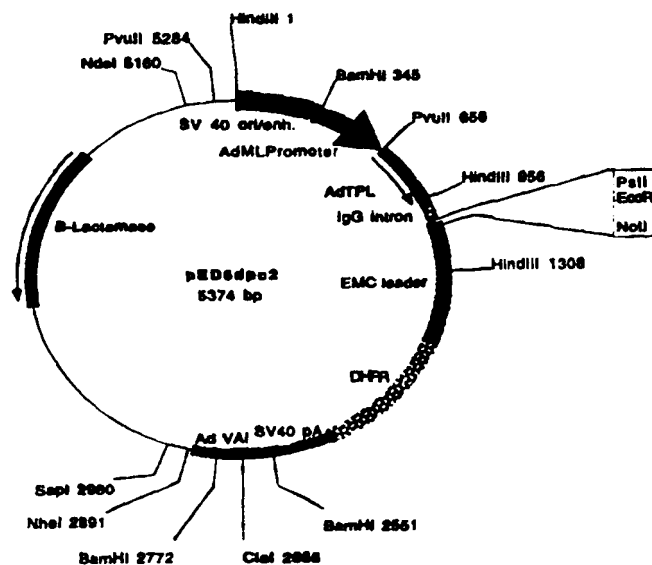
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 5/10, C07K 14/47, C12Q 1/68, A61K 38/17		A3	(11) International Publication Number: WO 98/32853
			(43) International Publication Date: 30 July 1998 (30.07.98)
(21) International Application Number: PCT/US98/01396		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 23 January 1998 (23.01.98)		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(30) Priority Data: 08/788,789 24 January 1997 (24.01.97) US		(88) Date of publication of the international search report: 7 January 1999 (07.01.99)	
(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).			
(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).			
(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).			

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kautman et al.(1991), NAR 19: 4485-4490.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

International Application No
PCT/98/01396

A. CLASSIFICATION OF SUBJECT MATTER IPC 6 C12N15/12 C12N5/10 C07K14/47 C12Q1/68 A61K38/17		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12N C07K C12Q A61K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	D. REISMAN ET AL.: "Human unknown protein mRNA within the p53 intron 1, complete cds." EMBL SEQUENCE DATABASE, 28 June 1996, HEIDELBERG, FRG, XP002074811 cited in the application Accession no. U58658 ---	1-13
A	L. HILLIER ET AL.: "The WashU-Merck EST Project" EMBL SEQUENCE DATABASE, 8 July 1995, HEIDELBERG, FRG, XP002074812 cited in the application yn72e01.r1 Homo sapiens cDNA clone 173976 5' similar to contains Alu repetitive element; Accession no. H23653; --- <div style="text-align: center;">-/--</div>	1-13
<div style="display: flex; justify-content: space-between;"> <input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex. </div>		
<div style="display: flex;"> <div style="flex: 1;"> <p>* Special categories of cited documents :</p> <p>*A* document defining the general state of the art which is not considered to be of particular relevance</p> <p>*E* earlier document but published on or after the international filing date</p> <p>*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>*O* document referring to an oral disclosure, use, exhibition or other means</p> <p>*P* document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="flex: 1;"> <p>*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>*Z* document member of the same patent family</p> </div> </div>		
Date of the actual completion of the international search <div style="text-align: center; font-size: 1.2em;">18 August 1998</div>		Date of mailing of the international search report <div style="text-align: center; font-size: 1.2em;">17. 11. 1998</div>
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer <div style="text-align: center; font-size: 1.2em;">HORNIG H.</div>

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/01396

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ADAMS M D ET AL: "3,400 NEW EXPRESSED SEQUENCE TAGS IDENTIFY DIVERSITY OF TRANSCRIPTS IN HUMAN BRAIN" NATURE GENETICS, vol. 4, no. 3, pages 256-267, XP000611495 see the whole document ---	1-13
A	JACOBS K ET AL: "A NOVEL METHOD FOR ISOLATING EUKARYOTIC CDNA CLONES ENCODING SECRETED PROTEINS" JOURNAL OF CELLULAR BIOCHEMISTRY - SUPPLEMENT, vol. 21A, 10 March 1995, page 19 XP002027246 see abstract ---	1-13
A	EP 0 510 691 A (OSAKA BIOSCIENCE INST) 28 October 1992 see the whole document ---	1-13
A	WO 94 07916 A (MERCK & CO INC ;SCHMIDT AZRIEL (US); RODAN GIDEON A (US); RUTLEDGE) 14 April 1994 see the whole document ---	1-13
A	WO 90 05780 A (OREGON STATE) 31 May 1990 see the whole document ---	1-13
A	WO 90 14432 A (GENETICS INST) 29 November 1990 see the whole document ---	1-13
A	WO 96 17925 A (IMMUNEX CORP) 13 June 1996 see the whole document ---	1-13
A	R.J. KAUFMAN ET AL.: "Effect of von Willebrand factor coexpression on the synthesis and secretion of factor VIII in chinese hamster ovary cells" MOL. CELL. BIOL., vol. 9, no. 3, March 1989, ASM WASHINGTON, DC,US, pages 1233-1242, XP002041592 see the whole document ---	1-13
A	R.J. KAUFMAN ET AL.: "The phosphorylation state of eucaryotic initiation factor 2 alters translation efficiency of specific mRNAs" MOL. CELL. BIOL., vol. 9, no. 3, March 1989, ASM WASHINGTON, DC,US, pages 946-958, XP002041593 see the whole document ---	1-13

-/--

INTERNATIONAL SEARCH REPORT

International Application No

PCT/98/01396

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	R.J. KAUFMAN ET AL.: "Improved vectors for stable expression of foreign genes in mammalian cells by use of the untranslated leader sequence from EMC virus" NUCLEIC ACIDS RESEARCH, vol. 19, no. 16, 1991, IRL PRESS LIMITED, OXFORD, ENGLAND, pages 4485-4490, XP002041594 cited in the application see the whole document ---	1-13
A	US 5 536 637 A (JACOBS KENNETH) 16 July 1996 cited in the application see the whole document ---	1-13
P,A	WO 97 07198 A (GENETICS INSTITUT) 27 February 1997 see the whole document ---	1-13
P,A	WO 97 25427 A (GENETICS INST) 17 July 1997 see the whole document -----	1-13

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 98/01396

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Please see Further Information sheet enclosed.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1 - 13

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-13

A composition comprising an isolated polynucleotide selected from the group consisting of: SEQ ID no.1; said composition wherein said polynucleotide is operably linked to an expression control sequence; a host cell transformed with said composition; a process for producing a protein which is encoded by said polynucleotide sequence; a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group of SEQ ID no.2, said composition further comprising a pharmaceutical acceptable carrier; a method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of said composition, the gene corresponding to the cDNA sequence of SEQ ID no.1;

2. Claims: 14-16

A composition comprising an isolated polynucleotide sequence selected from the group of SEQ ID no.3; a composition comprises a protein, wherein said protein comprises an amino acid sequence selected from the group of SEQ ID no.4; the gene corresponding to the cDNA sequences of SEQ ID nos.3;

3. Claims: 17-19

Idem as subject 2 but limited to SEQ ID nos.5 and 6;

4. Claims: 20-22

Idem as subject 2 but limited to SEQ ID nos.7 and 8;

5. Claims: 23-25

Idem as subject 2 but limited to SEQ ID nos. 9,10,11 and 12;

6. Claims: 26-28

Idem as subject 2 but limited to SEQ ID nos.13 and 14;

7. Claims: 29-31

Idem as subject 2 but limited to SEQ ID nos.15 and 16;

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

8. Claims: 32-34

Idem as subject 2 but limited to SEQ ID nos.17 and 18;

9. Claims: 35-37

Idem as subject 2 but limited to SEQ ID nos.19 and 20;

10. Claims: 38-40

Idem as subject 2 but limited to SEQ ID nos.21 and 22;

REMARK:

Although claim 12 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/JP98/01396

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0510691 A	28-10-92	CA 2067031 A JP 5184368 A	27-10-92 27-07-93
WO 9407916 A	14-04-94	AU 5165193 A	26-04-94
WO 9005780 A	31-05-90	AT 154636 T AU 645963 B AU 4668689 A DE 68928137 D DE 68928137 T EP 0447483 A ES 2104599 T JP 4506449 T	15-07-97 03-02-94 12-06-90 24-07-97 19-02-98 25-09-91 16-10-97 12-11-92
WO 9014432 A	29-11-90	US 5580753 A AT 147436 T AU 637620 B AU 5928990 A CA 2056997 A DE 69029657 D DK 473724 T EP 0473724 A ES 2099096 T JP 4506006 T US 5734037 A US 5414071 A	03-12-96 15-01-97 03-06-93 18-12-90 24-11-90 20-02-97 14-04-97 11-03-92 16-05-97 22-10-92 31-03-98 09-05-95
WO 9617925 A	13-06-96	AU 4639396 A CA 2206488 A EP 0871702 A FI 972390 A NO 972455 A NZ 301067 A	26-06-96 13-06-96 21-10-98 05-06-97 06-08-97 25-03-98
US 5536637 A	16-07-96	US 5712116 A	27-01-98
WO 9707198 A	27-02-97	US 5707829 A AU 6712396 A AU 6768596 A CA 2229208 A	13-01-98 18-02-97 12-03-97 27-02-97

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/01396

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9707198 A		EP 0839196 A	06-05-98
		EP 0851875 A	08-07-98
		WO 9704097 A	06-02-97
WO 9725427 A	17-07-97	AU 1532697 A	01-08-97

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 September 2000 (08.09.2000)

PCT

(10) International Publication Number
WO 00/52151 A3

(51) International Patent Classification⁷: **C12N 15/00,**
C07K 14/47, G01N 33/53

[US/US]: 55 Park Belmont Place, San Jose, CA 95136
(US). **AZIMZAI, Yalda** [US/US]; 2045 Rock Springs
Drive, Hayward, CA 94545 (US).

(21) International Application Number: **PCT/US00/05621**

(22) International Filing Date: **3 March 2000 (03.03.2000)**

(74) Agents: **HAMLET-COX, Diana et al.**; Incyte Pharma-
ceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304
(US).

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:
60/123,117 **5 March 1999 (05.03.1999) US**

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE,
ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP,
KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD,
MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD,
SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ,
VN, YU, ZA, ZW.

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier application:
US 60/123,117 (CIP)
Filed on **5 March 1999 (05.03.1999)**

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,
MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM,
GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **INCYTE
PHARMACEUTICALS, INC.** [US/US]; 3160 Porter
Drive, Palo Alto, CA 94304 (US).

Published:
— *With international search report.*

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **TANG, Y., Tom**
[CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US).
LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA
95054 (US). **BAUGHN, Mariah, R.** [US/US]; 14244
Santiago Road, San Leandro, CA 94577 (US). **YUE,**
Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087
(US). **AU-YOUNG, Janice** [US/US]; 233 Golden Eagle
Lane, Brisbane, CA 94005 (US). **LU, Dyung, Aina, M.**

(88) Date of publication of the international search report:
26 April 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/52151 A3

(54) Title: **HUMAN SECRETORY PROTEINS**

(57) Abstract: The invention provides human secretory proteins (HSECP) and polynucleotides which identify and encode HSECP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HSECP.



1

2

3

4

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/05621

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/00 C07K14/47 G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	WO 99 22243 A (ENDRESS GREGORY A ; FLORENCE KIMBERLY A (US); HUMAN GENOME SCIENCES) 6 May 1999 (1999-05-06) abstract page 101 -page 102 Seq.Id.No.215	1-23
A	SHIROZU M ET AL: "CHARACTERIZATION OF NOVEL SECRETED AND MEMBRANE PROTEINS ISOLATED BY THE SIGNAL SEQUENCE TRAP METHOD" GENOMICS,US,ACADEMIC PRESS, SAN DIEGO, vol. 37, no. 3, 1 November 1996 (1996-11-01), pages 273-280, XP002054773 ISSN: 0888-7543 the whole document	1-23

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

1 September 2000

Date of mailing of the international search report

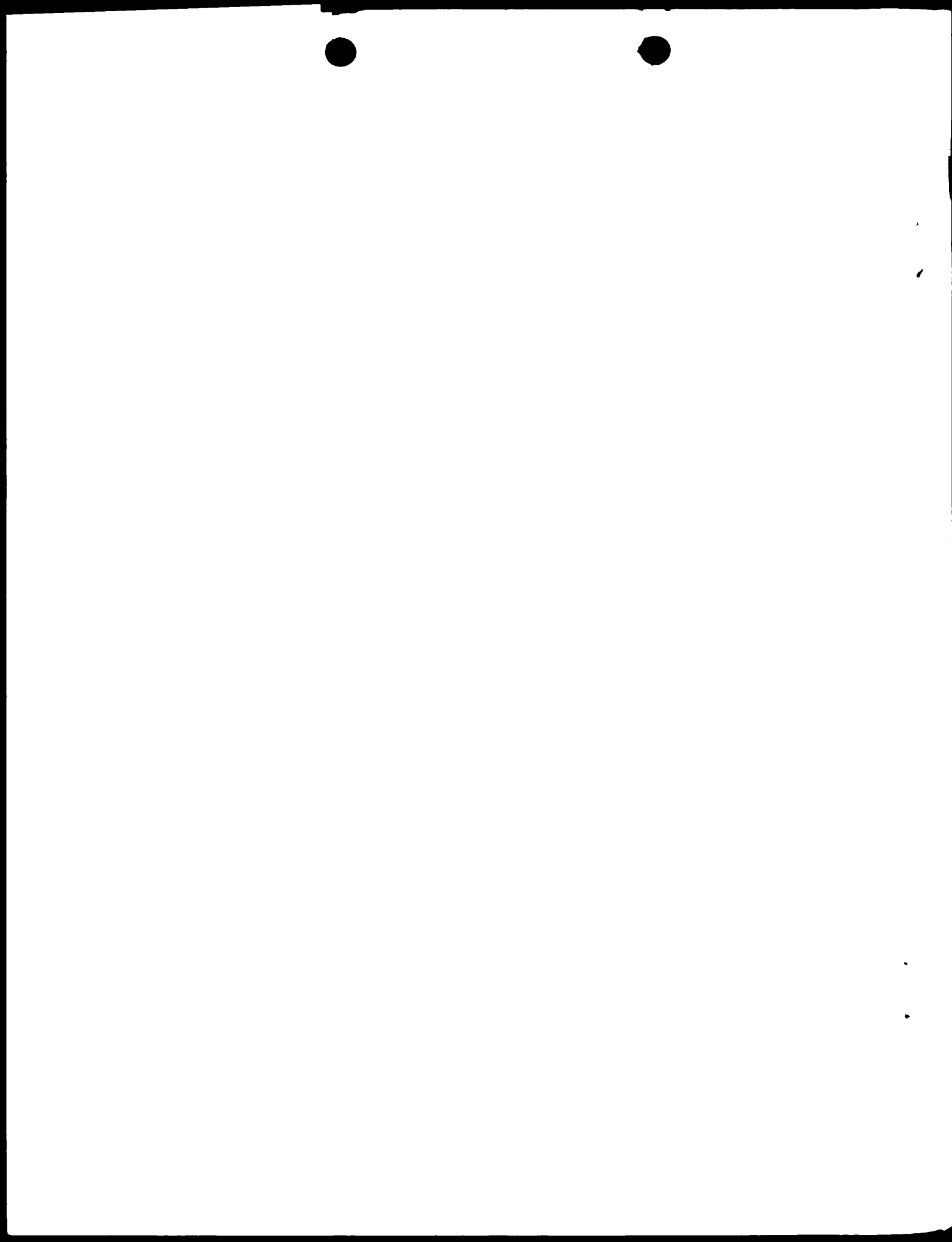
05. 1. 01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

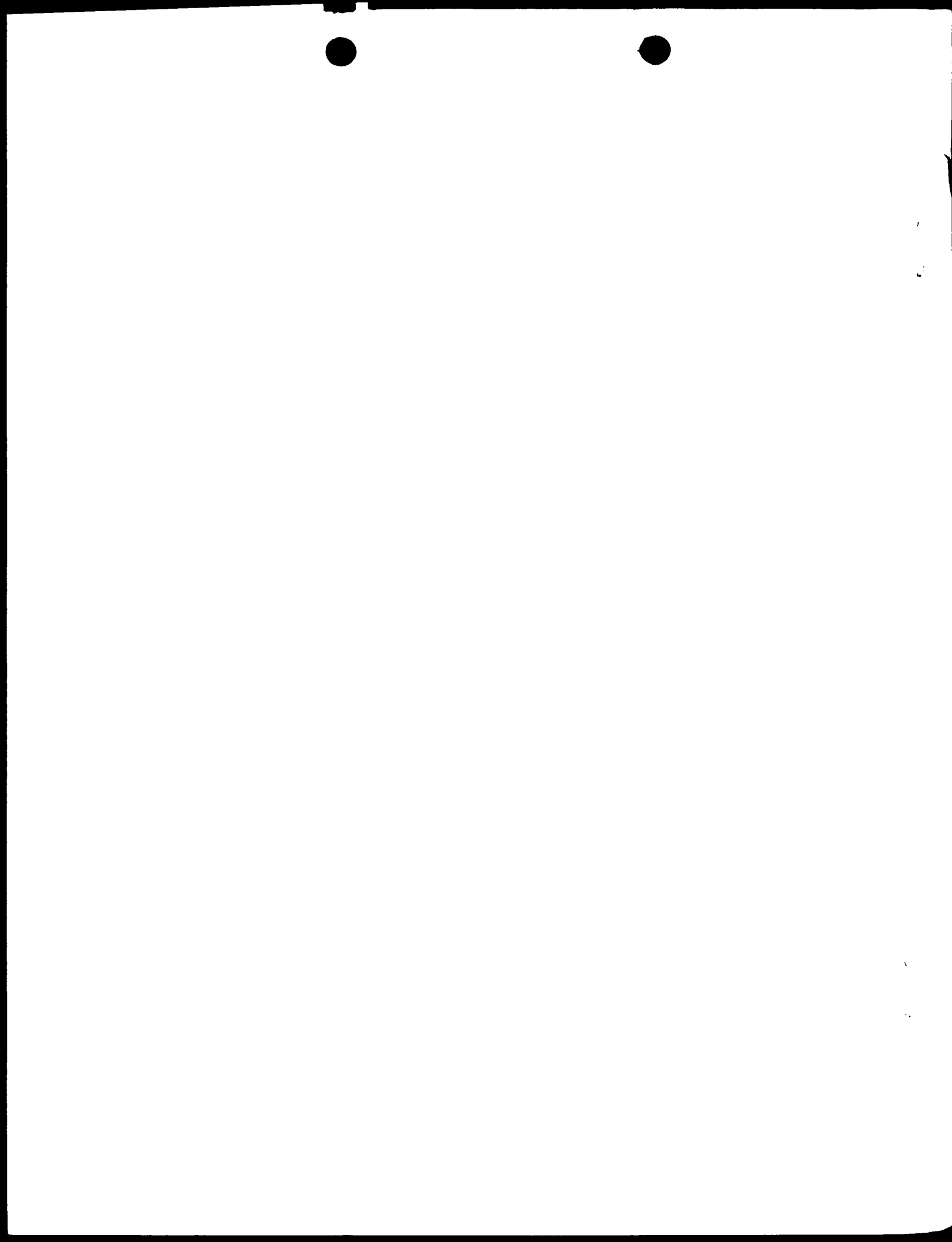
Panzica, G



INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/05621

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JACOBS K A ET AL: "A genetic selection for isolating cDNAs encoding secreted proteins" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES,GB,ELSEVIER SCIENCE PUBLISHERS, BARKING, vol. 198, no. 1-2, 1 October 1997 (1997-10-01), pages 289-296, XP004116069 ISSN: 0378-1119 the whole document	1-23
A	--- WO 98 30696 A (GENETICS INST) 16 July 1998 (1998-07-16) the whole document	1-23
A	--- WO 98 32853 A (GENETICS INST) 30 July 1998 (1998-07-30) abstract	1-23
A	--- WO 98 40486 A (GENETICS INST) 17 September 1998 (1998-09-17) abstract -----	1-23



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/05621

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 16, 19 and 22 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-23 (PARTIALLY)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.



FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-23 (partially)

An isolated polypeptide having sequence as set forth in Seq.Id.No.1 of the sequence listing and its nucleic acid sequence as set forth in Seq.Id.No.23. Methods of production and uses thereof.

2. Claims: 1-23 (partially)

Inventions 2-22:

Same as for invention no.1 but respectively to each following pair of aminoacid sequences with their respective nucleic acid sequences:

Invention 2: Seq.Id.Nos.2 and 24.

Invention 3: Seq.Id.Nos.3 and 25.

....

Invention 22: Seq.Id Nos.22 and 44.



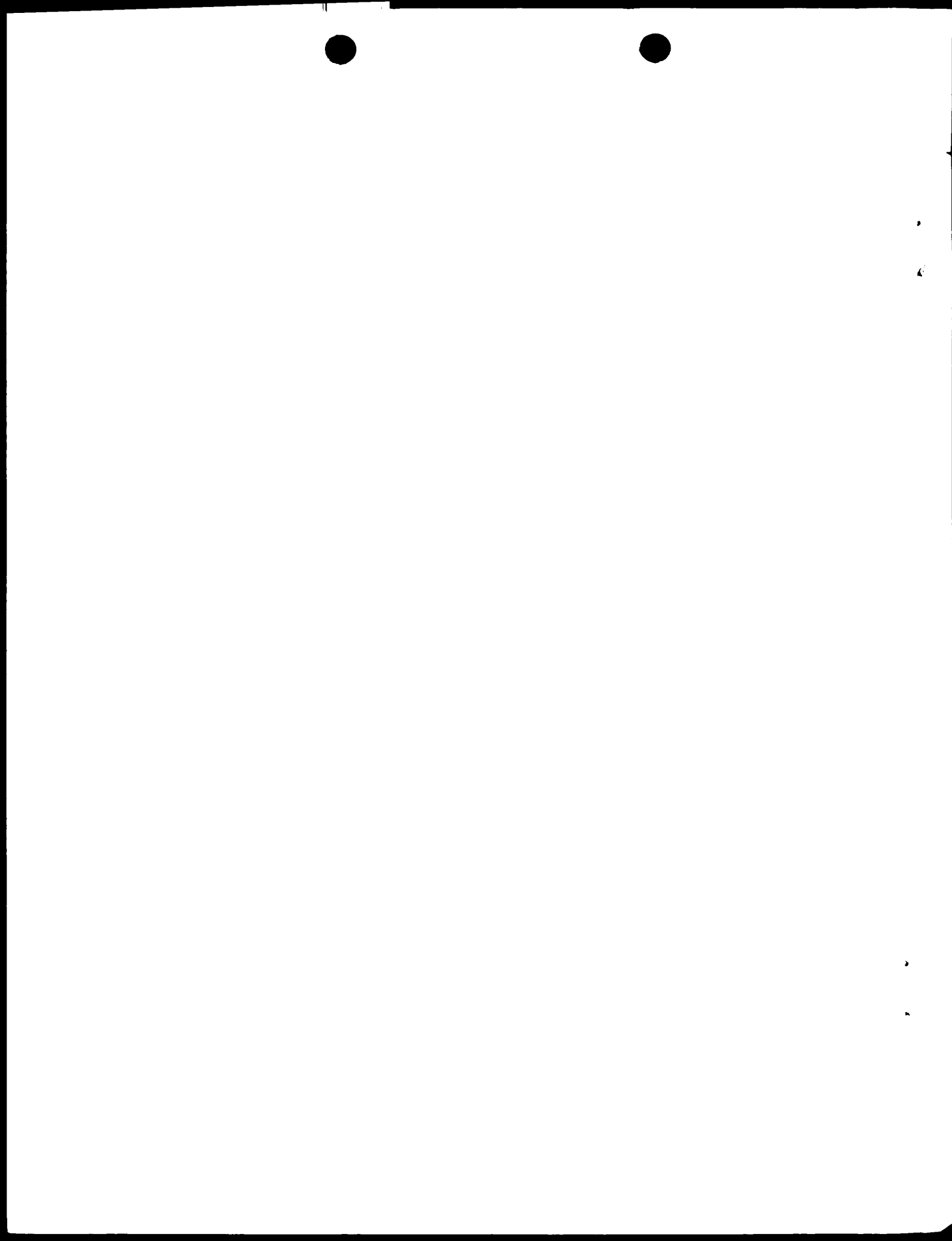
INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/05621

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9922243 A	06-05-1999	AU 1273499 A	17-05-1999
		EP 1042674 A	11-10-2000
		AU 1118499 A	17-05-1999
		WO 9921575 A	06-05-1999
WO 9830696 A	16-07-1998	US 5945302 A	31-08-1999
		US 5972652 A	26-10-1999
		AU 2723397 A	29-10-1997
		EP 0907732 A	14-04-1999
		EP 1007660 A	14-06-2000
		JP 2000508172 T	04-07-2000
		AU 5822398 A	03-08-1998
		AU 2558599 A	26-07-1999
		EP 1047713 A	02-11-2000
		WO 9935168 A	15-07-1999
WO 9832853 A	30-07-1998	AU 5828398 A	18-08-1998
		EP 1007661 A	14-06-2000
WO 9840486 A	17-09-1998	US 5976837 A	02-11-1999
		AU 6702298 A	29-09-1998
		EP 0973890 A	26-01-2000



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G01N 33/68, 33/53, C07K 16/00, C12N 15/11, 15/12, 15/00, 15/63, A61K 38/17, 38/16, C12P 21/02		A1	(11) International Publication Number: WO 99/22243																																																			
			(43) International Publication Date: 6 May 1999 (06.05.99)																																																			
(21) International Application Number: PCT/US98/22376																																																						
(22) International Filing Date: 23 October 1998 (23.10.98)																																																						
(30) Priority Data: <table border="0"><tr><td>60/063,099</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,088</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,100</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,387</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,148</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,386</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/062,784</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,091</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,090</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,089</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,092</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,111</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,101</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,109</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,110</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,098</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr><tr><td>60/063,097</td><td>24 October 1997 (24.10.97)</td><td>US</td></tr></table>		60/063,099	24 October 1997 (24.10.97)	US	60/063,088	24 October 1997 (24.10.97)	US	60/063,100	24 October 1997 (24.10.97)	US	60/063,387	24 October 1997 (24.10.97)	US	60/063,148	24 October 1997 (24.10.97)	US	60/063,386	24 October 1997 (24.10.97)	US	60/062,784	24 October 1997 (24.10.97)	US	60/063,091	24 October 1997 (24.10.97)	US	60/063,090	24 October 1997 (24.10.97)	US	60/063,089	24 October 1997 (24.10.97)	US	60/063,092	24 October 1997 (24.10.97)	US	60/063,111	24 October 1997 (24.10.97)	US	60/063,101	24 October 1997 (24.10.97)	US	60/063,109	24 October 1997 (24.10.97)	US	60/063,110	24 October 1997 (24.10.97)	US	60/063,098	24 October 1997 (24.10.97)	US	60/063,097	24 October 1997 (24.10.97)	US		
60/063,099	24 October 1997 (24.10.97)	US																																																				
60/063,088	24 October 1997 (24.10.97)	US																																																				
60/063,100	24 October 1997 (24.10.97)	US																																																				
60/063,387	24 October 1997 (24.10.97)	US																																																				
60/063,148	24 October 1997 (24.10.97)	US																																																				
60/063,386	24 October 1997 (24.10.97)	US																																																				
60/062,784	24 October 1997 (24.10.97)	US																																																				
60/063,091	24 October 1997 (24.10.97)	US																																																				
60/063,090	24 October 1997 (24.10.97)	US																																																				
60/063,089	24 October 1997 (24.10.97)	US																																																				
60/063,092	24 October 1997 (24.10.97)	US																																																				
60/063,111	24 October 1997 (24.10.97)	US																																																				
60/063,101	24 October 1997 (24.10.97)	US																																																				
60/063,109	24 October 1997 (24.10.97)	US																																																				
60/063,110	24 October 1997 (24.10.97)	US																																																				
60/063,098	24 October 1997 (24.10.97)	US																																																				
60/063,097	24 October 1997 (24.10.97)	US																																																				
(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).																																																						
(72) Inventors; and																																																						
(75) Inventors/Applicants (for US only): FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). WEI, Ying-Fei [CN/US]; 1714-C Marina Court, San Mateo, CA 94403 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). MOORE, Paul, A. [GB/US]; 19005 Leatherbark Drive, Germantown, MD 20874 (US). KAYW, Hia [MM/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US). LAFLEUR, David, W. [US/US]; 3142 Quesada Street, N.W., Washington, DC 20015 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick																																																						
		<p>Place #24, Gaithersburg, MD 29878 (US). BREWER, Laurie, A. [US/US]; Apartment 115, 410 Van Dyke Street, St. Paul, MN 55119-4321 (US). SHI, Yanggu [CN/US]; Apartment 102, 437 West Side Drive, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). YOUNG, Paul [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]; 872 Diamond Drive, Gaithersburg, MD 20878 (US). FLORENCE, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). FLORENCE, Charles [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). DUAN, D., Roxanne [US/US]; 5515 Northfield Road, Bethesda, MD 20817 (US). JANAT, Fouad [US/US]; 140 High Street #202, Westerly, RI 02891 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US).</p>																																																				
		(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).																																																				
		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).																																																				
		Published <i>With international search report.</i> <i>With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.</i>																																																				
(54) Title: 148 HUMAN SECRETED PROTEINS																																																						
(57) Abstract <p>The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.</p>																																																						

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

148 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and
5 their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or
10 organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum
15 (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

20 Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or
25 secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include
30 the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using
35 secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard,
5 Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained
10 in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42°C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the
15 filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages
20 of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even
25 lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include
30 Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such
35 as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

5 The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and
10 double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability
15 or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

 The polypeptide of the present invention can be composed of amino acids joined
20 to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs,
25 as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be
30 branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a
35 nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine.

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

25 Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

35 MRFISQQSCECVRPCMDVYVCVYISIHVYMDAHVYLCRICKTNMR (SEQ ID NO:309), RILRWVNCMACDLYLNKAVSVCAHVWMCMCVYISLYMYTWMP MCIYVEYVKQT (SEQ ID NO:310), NPENQLEISFPPRRQKMKLTLDLQVSQS SLVHSLLSDDFFSVSKEGCLWKPIILPSHFL (SEQ ID NO:311), LQTQISN YLMFVLHILHRYTWASMYTCIEIYTHYTSIHGRTHSQLC (SEQ ID NO:312).

IHMGIHVYMYRDIYTHIHHTWAHTLTALLRYKSHAIQLTHLNIR (SEQ ID NO:313), and/or MKWIFTVLILTSCFFTAGICEDGICSRIQL RDKIVQSAFRQ (SEQ ID NO:314). Polynucleotides encoding these polypeptides are also encompassed by the invention.

5 This gene is expressed primarily in neutrophils.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, particularly neutropenia and related conditions.

10 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and
15 wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 The tissue distribution in neutrophils indicates that polynucleotides and
20 polypeptides corresponding to this gene are useful for immune disorders. More specifically, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune
25 functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-
30 graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the
35 differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 812 of SEQ ID NO:11, b is an integer of 15 to 826, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

KPCCPSVSNRSSVQMHQLPIQLGQFEAHGFCRSFLETIFYTHDPRAMHSFL
SSISSPSLPFGFSRMTSQINHLHPSPLC (SEQ ID NO:315). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, particularly neutropenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:161 as residues: Asp-15 to Tyr-21, Pro-29 to Asn-39.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for immune disorders. Moreover,

the expression of this gene product indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness
5 in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease,
10 sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease,
15 scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

20 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
25 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 510 of SEQ ID NO:12. b is an integer of 15 to 524, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

In specific embodiments, polypeptides of the invention comprise the following
35 amino acid sequence: SVFKINLKSFKQHEPWWPNRS (SEQ ID NO:316). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, including neutropenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:162 as residues: Met-1 to Arg-8, Leu-35 to Glu-41.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for immune disorders. More specifically, expression of this gene product indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may also be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 477 of SEQ ID NO:13, b is an integer of 15 to 491, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

This gene is expressed primarily in IL-1 and LPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, including neutrophenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:163 as residues: Asn-45 to Thr-58.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or prevention of a variety of immune disorders. In particular, this gene product may play a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. Furthermore, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by

boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 389 of SEQ ID NO:14, b is an integer of 15 to 403, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

GTRSFSVPSYLRLTGSLMCYLLLLLIQTAELLIHPQGLQAVSNGESALKGTRPTF
SSPFILVTEGRKEWEGVFLSSGWKGNTLSNYYISLVFYYSRILQPYFYCLWGK
LEMVTLIRSVWRGINGGDKISVGFGKC (SEQ ID NO:317). WMERKHTVKLL
YLLGFLQLQNSPAIFLLSMGEVGDGDL (SEQ ID NO:318) SNGESALKGTRP
TFSSPFILVTE (SEQ ID NO:319), and/or LSNYYISLVFYYSRILQPYFYCLW (SEQ
ID NO:320). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome

17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

This gene is expressed primarily in the breast and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, reproductive, or neural disorders, such as cancers of the breast, lymph system and brain. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
10 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive, immune, and central nervous systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, reproductive, neural, breast, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, breast milk, serum, plasma, urine, synovial fluid and
15 spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:164 as residues: Leu-31 to Phe-38, Glu-47 to Trp-52.

20 The tissue distribution in breast and brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancers in the breast, lymph system, and brain. Moreover, the protein product of this gene may be useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease,
25 Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered
30 bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in
35 the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular

system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 799 of SEQ ID NO:15, b is an integer of 15 to 813, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 6

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, such as neutropenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:165 as residues: Ser-49 to Leu-54.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of a variety of immune disorders. Moreover, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product

may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also
5 used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-
10 host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of
15 various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present
20 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 250 of SEQ ID NO:16, b is an integer of 15 to
25 264, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

30

The translation product of this gene shares sequence homology with neurotoxin which is thought to be important in neural diseases. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:
EKDFMQGSDAGHGGTHIYRALVQWPLAWVFYLSHAKTHWGEELRFSFRRKN
35 LRLREAMRHETCQVTQLVA GKADSNLCLRDSETWFWPPLWAACSSLQATA
CRLSSPSKGLGASRECPWLASGRAALVSFL (SEQ ID NO:321). SLRVKGRKPR
LLYHSPARGTLWMLPGLCDCL ICRQWLVERSRLPRVGARTRFQSP SDTGWS

QLCQLPAV (SEQ ID NO:322), and/or ERSRLPRVGARTRFQSPSDTGWSQLC (SEQ ID NO:323). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in neutrophils.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and neural diseases, particularly neurodegenerative disorders, such as Alzheimers or Parkinson's. Similarly, polypeptides and antibodies directed to
10 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and neural systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hemaopoietic, neural, and cancerous and wounded tissues) or bodily
15 fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
20 NO:166 as residues: Gln-2 to Gly-10, Asp-77 to Phe-82.

The tissue distribution in neutrophils combined with the homology to the conserved neurotoxin protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for immune and neural diseases. Similarly, the protein product of this gene may be useful for the detection/treatment of
25 neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive
30 compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or
35 neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular

system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 506 of SEQ ID NO:17, b is an integer of 15 to 520, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 8

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: KHAFLMAHQFCVLSLAMQWSSCFQLVALPYLSL (SEQ ID NO:324). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, such as neutropenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of immune disorders. Furthermore, this gene product may be involved in the

regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 979 of SEQ ID NO:18, b is an integer of 15 to 993, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

30

When tested against Jurket and PC12 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) and EGR1 (early growth response gene 1) promoter elements. Thus, it is likely that this gene activates T-cells and sensory neurons through the JAK-STAT and EGR1 signal transduction pathways. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.

Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon
5 activation, leading the cells to undergo differentiation and proliferation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: MRPLCVLLPWCWQWGGLGSASPIRPQAPPGQAAHAVP LPRAQHLAQRSRQ (SEQ ID NO:325). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is
10 believed to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

This gene is expressed primarily in breast, lymph nodes, spleen, and to a lesser extent, in liver.

Therefore, polynucleotides and polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, immune, or hepatic disorders, particularly cancers of the breast, liver, and lymph system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
20 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast, liver and lymph system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. reproductive, breast, immune hematopoietic, hepatic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, breast milk, bile, serum, plasma, urine,
25 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
30 NO:168 as residues: Pro-54 to Gly-67.

The tissue distribution in breast and immune tissues combined with the detected EGR1 and GAS biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancers of the breast and lymph systems. Moreover, the GAS and EGR1 activity strongly indicates
35 that the protein product of this gene may play an integral role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, such proliferative tissues rely on finely regulated

decisions involving cell differentiation and/or apoptosis. Thus this protein may also be involved in regulating apoptosis or tissue differentiation and, thus could be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- 10 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 445 of SEQ ID NO:19, b is an integer of 15 to 459, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

- In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ARGLRSPHGAAGVVRGDGGGKKGEDPYSPILFQ
20 SERIPRLIYLPVISSEENS (SEQ ID NO:326). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in an LPS induced neutrophil cDNA library.

- Therefore, polynucleotides and polypeptides of the invention are useful as
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system disorders, such as neutropenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of
30 disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the
35 standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders, for example, in ameliorating an aberrant neutrophil response to infectious agents. Similarly, the expression of this gene product may suggest a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may also be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 541 of SEQ ID NO:20, b is an integer of 15 to 555, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

This gene is expressed primarily in prostate cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, reproductive or immune system disorders, particularly prostate cancer.
Similarly, polypeptides and antibodies directed to these polypeptides are useful in
providing immunological probes for differential identification of the tissue(s) or cell
10 type(s). For a number of disorders of the above tissues or cells, particularly of the
immune system, expression of this gene at significantly higher or lower levels may be
detected in certain tissues or cell types (e.g. reproductive, prostate, and cancerous and
wounded tissues) or bodily fluids (e.g. lymph, seminal fluid, serum, plasma, urine,
synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual
15 having such a disorder, relative to the standard gene expression level, i.e., the
expression level in healthy tissue or bodily fluid from an individual not having the
disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
NO:170 as residues: Pro-14 to Asp-25, Leu-51 to Val-63.

20 The tissue distribution in prostate tissues indicates that polynucleotides and
polypeptides corresponding to this gene are useful for the diagnosis and treatment of
reproductive system disorders such as cancer, particularly prostate cancer. Similarly,
the expression within prostate cancer tissue, a cellular source marked by proliferating
cells, indicates that this protein may play a role in the regulation of cellular division, and
25 may show utility in the diagnosis and treatment of cancer and other proliferative
disorders not limited to prostate tissue. Further, such tissues rely on decisions
involving cell differentiation and/or apoptosis. Thus this protein may also be involved
in apoptosis or tissue differentiation and could again be useful in cancer therapy.
Protein, as well as, antibodies directed against the protein may show utility as a tumor
30 marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
and accessible through sequence databases. Some of these sequences are related to SEQ
ID NO:21 and may have been publicly available prior to conception of the present
invention. Preferably, such related polynucleotides are specifically excluded from the
35 scope of the present invention. To list every related sequence is cumbersome.
Accordingly, preferably excluded from the present invention are one or more
polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 651 of SEQ ID NO:21, b is an integer of 15 to 665, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

The polynucleotide sequence of this gene may have a frame shift. Therefore the preferred signal peptide may reside in a frame other than the associated polynucleotides of this gene. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

KSLSCSFLFLAFWLRRMGQTMCVVCVCVCVCVVRTWVYLYEPVKF RSPLIYV
NLPTS (SEQ ID NO:327), and/or KLGFTMLARLVSNSXTSGDLPSSASQNAGI
KGMSYRAWPYSYFLIRKNKQT NKQTKTNPQLGENKHCRNLKVSWSKNYFL
(SEQ ID NO:328). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system disorders, particularly immunodeficiencies such as lupus and AIDS, or inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Moreover, this gene product may play a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may

also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 763 of SEQ ID NO:22, b is an integer of 15 to 777, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

When tested against Jurket and U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates T-cells and promyelocytic cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the

proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

ERGQGGSSRNVAGSDLVFPVAVFVSXLC (SEQ ID NO:329),

GSPQGPSVALGSRQCWSRPLRRGGRGA AVEMWRGPTWCFRPSLCLCCVCGV

5 SFGLYVPHGFSLSMCVSAP GSAWLSLVYSICLARGSMSXRXSSRXSLV

ASGASVLLVCFWVXADPGVGVSVPRAXVSGLWWCVSPSACLXLAPTKPPP

XLSFSLSIFPFSSNPSK (SEQ ID NO:330), and/or TIASLQPTALNHLIWRGW

KRKGRLRERKRGXGGAWLGPXRGRQMDSHTTRDQRQXLGEQRHPLLGLXA

PRSKPTKQMPQMOPGXPEKKXXLTWNHGLDRWNTQGTARQSLGQK

10 HTWRD (SEQ ID NO:331). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in adipose tissue and the brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
15 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic or neural disorders, particularly obesity, and neurodegenerative or central nervous system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders
20 of the above tissues or cells, particularly of the brain and central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. adipose, neural, immune, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to
25 the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in adipose and neural tissues, combined with the detected GAS biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of obesity and disorders of the brain
30 and central nervous system. Similarly, the protein product of this gene may be useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal
35 cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in

feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. In addition, the protein product of this gene may also be beneficial in detecting, treating, or preventing neural disorders which occur secondary to aberrant fatty acid metabolism in neural tissues, such as for aberrations in myelin sheath development, or associated autoimmune disorders of neural tissue or the overlying integument. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 526 of SEQ ID NO:23, b is an integer of 15 to 540, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ARGPGTEGCEPWLQLQDRRER (SEQ ID NO:332), and/or MSSGTNSFFTLMALNSPTGDSGSRITVSPPRVHPVKSGRGRASDLLLTRFLAPR SALWS (SEQ ID NO:333). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in a cDNA library from IL-1 and LPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, immune system disorders, such as neutropenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of
5 this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
10 individual not having the disorder.

The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of conditions where lymphocytes show aberrant response to an infectious agent. Similarly, this gene product may play a role in regulating the proliferation; survival;
15 differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune
20 functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-
25 graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the
30 differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
35 ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 470 of SEQ ID NO:24, b is an integer of 15 to 484, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

10 This gene is expressed primarily in ovaries, tonsils, and CD34 positive bone marrow stem cells.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, immune, developmental, or hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and reproductive systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. reproductive, ovarian, immune, tonsil, umbilical, developmental, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in ovarian and tonsil tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immune and reproductive system disorders. Similarly, expression of this gene product in tonsils indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia.

neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 693 of SEQ ID NO:25, b is an integer of 15 to 707, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates promyelocytic cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.

Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: HEYHLLSSRHILGSVLRDVC SALWS (SEQ ID NO:334). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in IL-1 and LPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, such as neutropenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune cells combined with the detected GAS biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of immune disorders. Specifically, the expression of this gene product in neutrophils indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the

scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 779 of SEQ ID NO:26, b is an integer of 15 to 793, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

10

This gene is expressed primarily in the spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, any of a variety of nervous system and neuromuscular disorders, particularly amyotrophic lateral sclerosis, muscular dystrophy, and inherited and non-inherited forms of chorea. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and neuromuscular systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, neuromuscular, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spinal cord tissue indicates that this gene could be used for the treatment of spinal cord and related injuries. The protein product of this gene could be injected into the spinal cord to promote or control growth following injuring or degeneration. Alternatively cells expressing this gene could be injected or transferred into the spinal cord by other means as a treatment promoting the regulation of growth following spinal cord injury or degeneration. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction.

aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 624 of SEQ ID NO:27. b is an integer of 15 to 638, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

25

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates promyelocytic cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

35

FILFILEYDMLWKSLEYTNSSAYGYVIASYFCLLGKLLVKQKKKKKTRGGAR

X PIRPXVESYYKSXAVVLQRRGLGKNLGG (SEQ ID NO:335). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in the adrenal gland.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endocrine disorders, particularly disorders of the adrenal gland. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
10 type(s). For a number of disorders of the above tissues or cells, particularly of the adrenal gland, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. endocrine, adrenal, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a
15 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in adrenal tissue, combined with the detected GAS biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine
20 disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothalamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a
25 tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the
30 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 514 of SEQ ID NO:28, b is an integer of 15 to 528, where both a and b correspond to the positions of nucleotide residues shown in
35 SEQ ID NO:28, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

This gene is expressed primarily in the placenta.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
10 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. placental, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or
15 cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:178 as residues: His-15 to Trp-20, Pro-48 to Ala-54.

20 The tissue distribution in placental tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of reproductive disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

25 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

30 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 905 of SEQ ID NO:29, b is an integer of 15 to 919, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

The translation product of this gene shares sequence homology with human erythrocyte membrane anion-transport protein which is thought to be important in autoimmune diseases. Furthermore, the translation product of this gene also has homology to a human gnlPID1026838 (AB012130) sodium bicarbonate cotransporter2 which is thought to be important in maintaining cellular homeostasis. Contact of cells with supernatant expressing the product of this gene was found to increase the permeability of the plasma membrane of enterocytes and renal mesangial cells to calcium. Thus it is likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product binds a receptor on the surface of the enterocytes and renal mesangial cells. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating cellular processes within enterocytes and renal mesangial cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

RVSSHLFRLFGGLILDIKRKAPFFLSDFKDALSLQCLASILFLYCACMSPVITFG
GGLGEATEG RIVSTKIGSGQAFSSSEASVCMHLSHYSYFYLKSLPTA (SEQ ID
NO:336), FRLFGGLILDIKRKAPFFLSDFKD (SEQ ID NO:337), FLYCACMSPV
ITFGGGLGEATEG (SEQ ID NO:338), and/or SSSEASVCMHLSHYSYFYLKSL
(SEQ ID NO:339). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in human testes tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or reproductive disorders, particularly autoimmune diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, reproductive, testicular, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression

level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in testis, the homology to an erythrocyte membrane anion-transport protein, in addition to, the detected calcium flux biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment of autoimmune diseases and other immune diseases such as cancer, particularly in, but not limited to, testicular tissue. Similarly, the translation product of this gene may be important in maintaining normal cellular homeostasis. Therefore, the protein, as well as antibodies directed to the invention, is beneficial as a therapeutic in order to ameliorate conditions related to aberrant cellular pH regulation (for example, use antibodies to decrease the presence of the protein, or possibly in gene therapy applications in order to replace a defective form, or alternatively, increase the expression of either the endogenous or modified form of the invention). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 850 of SEQ ID NO:30, b is an integer of 15 to 864, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene is expressed primarily in the brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural disorders, which include, but are not limited to, disorders of the brain and central nervous system, such as neurodegenerative conditions and/or depression. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s)

or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:180 as residues: His-13 to Leu-18.

The tissue distribution in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of disorders of the brain and central nervous system. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 905 of SEQ ID NO:31, b is an integer of 15 to

919, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 22

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

PCLQVIGIDFCRLLLMCLVLKRNLTVPFSSYSPLKTITCITSEQIAVVSNNFFRQKL
10 GVRAK FFQGACLHTSKVVICLNLPISIQRADIRMWWLVVNTPYARGVNN
(SEQ ID NO:340). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in the spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, any of a variety of nervous system and neuromuscular disorders including, but not limited to, amyotrophic lateral sclerosis, muscular dystrophy, and inherited and non-inherited forms of chorea. Similarly, polypeptides and antibodies
20 directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and neuromuscular systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, cancerous and wounded tissues) or bodily
25 fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spinal cord indicates that this gene could be used for
30 the treatment of spinal cord injuries. The protein product of this gene could be injected into the spinal cord to promote or control growth following injuring or degeneration. Alternatively cells expressing this gene could be injected or transferred into the spinal cord by other means as a treatment promoting the growth or regulation of growth following spinal cord injury or degeneration. This gene may also be useful for the
35 detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases.

peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 942 of SEQ ID NO:32, b is an integer of 15 to 956, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: VVSVCVLETGQLGPAALCRSV (SEQ ID NO:341).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, which include, but are not limited to, inflammatory diseases or neutropenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification

of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of the inflammatory conditions. Moreover, This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 552 of SEQ ID NO:33, b is an integer of 15 to 566, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

5 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

NISVHGFPVPCRLQRLQGPCHPKCCPHXISSGKPRSSFSPSSYHCKFSRNATLL
VVPNIFSYMQSSFLIPQTSKYIILXPYAXTXRPIKXIFKQAKQ (SEQ ID NO:342),
IYNDMMMEKKKTEVYQKRXSGDNTWGGKGLVAFVSSMEQGIHVQRCFIANL
10 KFSSPGV (SEQ ID NO: 343), YDDGEKEDRGLPEEMXWGQHLGWQGPCSL
CLKHGTGNPCTEMFYCQFKIFISWCLIPLVFARLGDFRDRPGWIFSWRYHLKH
TVWGGYNIIML (SEQ ID NO: 344), TPGDENFKLAIKHLCTWIPCS (SEQ ID NO:
345), IRHEIFLTIESFCPSAPRGEDDDNLLRTSRVPDI (SEQ ID NO:346), IRGSIP
GHKKMHLS FNVAQAQWSLLKPLVLREEGALFLTHDQLESKNSWTLSIGPRV
15 PYTYVVVTWSSALWDLPNQPLAGRKESGGSYGPISVTQSPHQAALKWFAKK
KGKQSHSTVQLANILHVFXAPDXYHFVNTSLQLFLEYTVMCMCLENK QKT
LGR (SEQ ID NO:347), EPEVTQVXSXELTFQ PRKAGAKVTAGKSHHQVIHWE
FEIMLSSYSTDVPLWFLKFFSSNLPQTYFPHSGVKKWGSCFSLPWRDSPPLT
FISLLSSHLTTFHLYHLHHGIICLGFSVYFHRA YTSLCILETAVGSY (SEQ ID
20 NO:348), WSLKPLVLREEGALFLTHDQLESK (SEQ ID NO:349), WFAKKK
GKQSHSTVQLANILHV (SEQ ID NO:350), AGKSHHQVIHWEFEIMLSSYSTDVP
(SEQ ID NO:351), and/or HGIICLGFSVYFHRA YTSLCILETAV (SEQ ID NO:352).
Polynucleotides encoding these polypeptides are also encompassed by the invention.

 This gene is expressed primarily in smooth muscle.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, muscular or vascular disorders, which include, but are not limited to defective organ innervation; deficiencies in neuronal survival; peristaltic abnormalities;
30 digestive disorders; perturbations of the vasculature. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the smooth muscle, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types
35 (e.g., cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in smooth muscle tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders that result from failures of normal smooth muscle function. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it may be involved in controlling the digestive process, and such actions as peristalsis. Similarly, it may be involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1550 of SEQ ID NO:34, b is an integer of 15 to 1564, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 25

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: KRLTINARVHLWTLKSVPL (SEQ ID NO:353), EYVFNMX XYSKSRAISPLSGPYTPRGTTPLPIPEPGARQRDHPAS LKYAKIIQTKLFAL PYPKETSMKAVA (SEQ ID NO:354), and/or ETVPPRSSQFLKITXGPARSMSLIX XAIQNPEPYLLYLALIPQEALLLYLSSQSQVPGNETTPPV (SEQ ID NO:355). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system disorders, which include, but are not limited to lupus.

inflammatory conditions, and immunodeficiencies such as AIDS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:184 as residues: Ser-21 to Thr-34, Thr-38 to Glu-43.

The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in T-cells indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the

scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1021 of SEQ ID NO:35, b is an integer of 15
 5 to 1035, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 26

10

The translation product of this gene was determined to have homology to the human IB3089A protein which is thought to play an important role in tumor suppression (See Genbank Accession No.gil3041877 (AF027734)).In specific
 15 embodiments, polypeptides of the invention comprise the following amino acid sequence: NEVSFSLSLGFSPREFARWKVNNLAL ERKDFFSLLPLAPEFIRNI RLLGRRPNLQQVTENLIKKGTHFLLSATLGGKQHHNPKLIGCQTIGNNV KTRVA (SEQ ID NO:356), VPYFLIRFSVTCCRLGLLPRRRMFRIN SGARGNG KLKKSFLSRAK LFTFQRANSLGEKPRDKEKLTFSQSKRHKI (SEQ ID NO:357), and/or EMSAVLFNQIFCNLLQIGSPSKEANVPDKLWGKRQWQTEEVLPFQSQV
 20 VHLPTGKLPGGKAKG (SEQ ID NO:358). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in human fibrosarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
 25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders afflicting endothelial, muscular, and extracellular matrix tissues, which include, but are not limited to fibrosarcomas and bladder cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
 30 type(s). For a number of disorders of the above tissues or cells, particularly of the integumentary system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g.endothelial, urogenital,renal, muscular, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken
 35 from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:185 as residues: Pro-49 to Asp-68.

The tissue distribution in human fibrosarcoma indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of various cancers, particularly fibrosarcomas and fibroids. Moreover, the expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 606 of SEQ ID NO:36, b is an integer of 15 to 620, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

This gene is expressed primarily in human tonsil.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, which include, but are not limited to inflammation and infectious diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types

(e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in tonsils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis of inflammation and infectious diseases. Moreover, this gene product may play a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 959 of SEQ ID NO:37, b is an integer of 15 to 973, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

5 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

HYHGSGLIKEFGSFLSLLCMLSCPYVFCHGMLEQEVPSVSPSTLDF PTSR
TVNKFLFKLPSLWYSVIATQNGLKQKIRETFLFVQFSQMPRWHKLE (SEQ ID
NO:359). Polynucleotides encoding these polypeptides are also encompassed by the
10 invention.

This gene is expressed primarily in adipose and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, metabolic or neural conditions, which include but are not limited to obesity and disorders of the brain and central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system,
20 expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, metabolic tissues, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily
25 fluid from an individual not having the disorder.

The tissue distribution in neural and adipose tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of obesity and disorders of brain and central system. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of
30 neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive
35 compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain

indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. In addition, considering the expression within both adipose tissue and brain indicates that the protein may be beneficial either as a target for gene therapy, or as a novel therapeutic to ameliorate conditions affecting myelin sheath development in neurons, or other disorders involving neural tissue which occur secondary to aberrant fatty-acid metabolism. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 824 of SEQ ID NO:38, b is an integer of 15 to 838, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 29

The gene encoding the disclosed cDNA is thought to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11. One embodiment of this gene comprises polypeptides of the following amino acid sequence:

FCKHNGSKNVFSTFRTPAVLFTGIVALYIASGLTGFIGLEVVAQLFNC (SEQ ID NO:360). An additional embodiment is the polynucleotides encoding the polypeptides.

This gene is expressed primarily in suppressor T cells, endothelial cells, dendritic cells, and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, immune system disorders related to abnormal activation of T cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. hematopoietic, developmental, neural, immune, endothelial, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:188 as residues: Tyr-14 to Leu-24. Pro-59 to Gln-66.

The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating disorders of the immune system related to altered activation of T cells. Furthermore, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of immune disorders. Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 593 of SEQ ID NO:39, b is an integer of 15 to

607, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 30

This gene is expressed primarily in the fetus and in tumor cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of rapidly growing tissues such as cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of rapidly growing tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. fetal, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene primarily in the developing fetus indicates a role in the treatment and/or detection of developmental disorders and growth defects. In addition, expression in tumor cell types indicates a role in the detection and/or treatment of tumors. Furthermore, expression within fetal tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 868 of SEQ ID NO:40, b is an integer of 15 to 882, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

This gene is expressed primarily in salivary gland, and to a lesser extent, in
10 other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, digestive and immune disorders. Similarly, polypeptides and antibodies
15 directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the salivary gland and other glands of the exocrine system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. exocrine, digestive, cancerous and
20 wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
25 NO:190 as residues: Glu-25 to Gly-31, Tyr-62 to Thr-68.

The tissue distribution in salivary gland tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of digestive and immune system disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the
30 above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the
35 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 945 of SEQ ID NO:41, b is an integer of 15 to 959, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

The gene encoding the disclosed cDNA is thought to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 12.

This gene is expressed primarily in brain tissue of adults, as well as infants.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative and behavioural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central and peripheral nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. brain, developmental, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:191 as residues: Ser-16 to Val-33.

The tissue distribution in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntintons Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Furthermore, expression of this gene product within the brain indicates that it may be involved in neuronal survival; synapse formation; conductance; neural differentiation, etc. Such involvement may impact many processes, such as learning and cognition. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 861 of SEQ ID NO:42, b is an integer of 15 to 875, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

This gene is expressed primarily in the synovium.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases affecting the synovial lining including arthritis and autoimmune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the musculo-skeletal system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. endothelial, skeletal, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for use as a factor that may protect against articular damage or promote growth of the cells in articulating joints. Furthermore, the expression of this gene product in synovium would suggest a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and

dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 616 of SEQ ID NO:43, b is an integer of 15 to 630, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

20

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

30

This gene is expressed primarily in B-cell lymphoma cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, such as diseases of B-cell lineage including lymphomas lymphoblastic leukemias, myelomas and hairy cell leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a

number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:193 as residues: Lys-82 to Pro-90.

The tissue distribution and biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and or diagnosis of diseases of B-cell lineage including cancer. This factor may be useful in the terminal differentiation of malignant cells or may act as a growth factor for B-cell proliferation or differentiation, which is supported by the biological assay data. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 557 of SEQ ID NO:44, b is an integer of 15 to 571, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS

element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in osteoclastoma derived stromal cells, placenta, pancreas and several tumor derived cells and to a lesser extent in brain, melanocytes, dendritic cells, and several other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of the pancreas, uterus, ovary, bone, or adrenal gland. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. placenta, pancreas, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating or diagnosing tumors of the reproductive organs, pancreas, or bone marrow. Furthermore, polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothalamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 916 of SEQ ID NO:45, b is an integer of 15 to

930, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 36

When tested against K562 leukemia cell lines, supernatants removed from cells containing this gene activated the ISRE assay. Thus, it is likely that this gene activates leukemia cells through the Jak-STAT signal transduction pathway. The interferon-sensitive response element is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in kidney and to a lesser extent in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal and nervous system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal and nervous systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. renal, urogenital, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:195 as residues: Lys-117 to Lys-126.

The tissue distribution of this gene in kidney tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or detection of renal disorders including kidney failure and Wilms Tumor in addition to the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntintons Disease,

schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 423 of SEQ ID NO:46, b is an integer of 15 to 437, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 37

One embodiment of this gene comprises polypeptides of the following amino acid sequence:

MPKPGAATQRTLLCLPRLHPASGPPLPXAGPLRGLRQLPALVPAASCRRRPAP
20 RLCAAGPCTVGPAASPHAPPHGCPPPASLAHV AHRQSVSGTVCLGLRDGHV
RGGCAAVRGXAALPWDAAAAGPDHMGVVGSGPALL (SEQ ID NO:361). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in pituitary and to a lesser extent in thymus and breast.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types or cell types (e.g. thymus, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or
35 another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or detection of endocrine, metabolic, and immune disorders including growth and developmental defects, in addition to the treatment or detection of immune or hematopoietic disorders including arthritis, asthma, immunodeficiency diseases and leukemia. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1010 of SEQ ID NO:47, b is an integer of 15 to 1024, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

20 **FEATURES OF PROTEIN ENCODED BY GENE NO: 38**

This gene is expressed primarily in hemangiopericytoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, vascular disorders such as stroke, aneurysm, cardiac arrest, hemorrhage. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. circulatory system, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:197 as residues: Cys-14 to Gly-23, Met-45 to Gly-51.

The tissue distribution of this gene solely in hemangiopericytoma indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and or detection of vascular disorders including hemorrhaging, aneurysm, stroke and cardiac arrest. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 449 of SEQ ID NO:48, b is an integer of 15 to 463, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

The translation product of this gene shares sequence homology with a serine protease which is thought to be important in regulating the availability and action of proteins in vivo.

This gene is expressed primarily in cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system related to abnormal growth factor regulation, including neurodegenerative conditions such as Alzheimers disease and psychiatric illness such as Schizophrenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the Central Nervous System, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. CNS, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression

level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:198 as residues: Ser-17 to Gln-22.

5 The tissue distribution in neural tissue, combined with the homology to serine proteases indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating disorders of the central nervous system including neurodegenerative diseases and psychiatric disorders. Furthermore, expression of this gene product within cerebral tissue indicates that it may be involved in neuronal survival; synapse formation;
10 conductance; neural differentiation, etc. Such involvement may impact many processes, such as learning and cognition. It may also be useful in the treatment of such neurodegenerative disorders as schizophrenia; ALS; or Alzheimer's. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
20 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 871 of SEQ ID NO:49, b is an integer of 15 to 885, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

This gene is expressed primarily in CD34 depleted buffy coat.

30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, autoimmune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
35 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g.

immune, developmental, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in CD34 depleted buffy coat tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating disorders of the immune system including autoimmune diseases. Furthermore, expression of this gene product in CD34 depleted buffy coat indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 833 of SEQ ID NO:50, b is an integer of 15 to 847, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

This gene is expressed primarily in B-cell lymphoma cells.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of B-cell lineage including lymphomas lymphoblastic leukemias, myelomas and hairy cell leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
10 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
15 individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for for the treatment and or diagnosis
20 of diseases of B-cell lineage including cancer. This factor may be useful in the terminal differentiation of malignant cells or may act as a growth factor for B-cell proliferation or differentiation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
25 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more
30 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 566 of SEQ ID NO:51, b is an integer of 15 to 580, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

This gene is expressed primarily in brain and CD34 depleted buffy coat.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, autoimmune disorders particularly those of the central nervous system
such as multiple sclerosis. Similarly, polypeptides and antibodies directed to these
polypeptides are useful in providing immunological probes for differential identification
10 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
particularly of the immune system, expression of this gene at significantly higher or
lower levels may be detected in certain tissues or cell types (e.g. immune, neural, and
cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum,
plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken
15 from an individual having such a disorder, relative to the standard gene expression
level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
NO:201 as residues: Pro-35 to Ala-40.

20 The tissue distribution indicates that polynucleotides and polypeptides
corresponding to this gene are useful for treating autoimmune disorders such as
multiple sclerosis. Furthermore, expression of this gene product in CD34 depleted
buffy coat indicates a role in the regulation of the proliferation; survival; differentiation;
and/or activation of potentially all hematopoietic cell lineages, including blood stem
25 cells. This gene product may be involved in the regulation of cytokine production,
antigen presentation, or other processes that may also suggest a usefulness in the
treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed
in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against
the protein may show utility as a tumor marker and/or immunotherapy targets for the
30 above listed tissues. Therefore it may be also used as an agent for immunological
disorders including arthritis, asthma, immune deficiency diseases such as AIDS,
leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and
psoriasis. In addition, this gene product may have commercial utility in the expansion
of stem cells and committed progenitors of various blood lineages, and in the
35 differentiation and/or proliferation of various cell types. Protein, as well as, antibodies
directed against the protein may show utility as a tumor marker and/or immunotherapy
targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 584 of SEQ ID NO:52, b is an integer of 15 to 598, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

This gene is expressed primarily in tissues of the brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and neurodegenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. CNS, brain, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful as a neuronal protective agent and as a growth factor for cells of the central or peripheral nervous system. Furthermore, expression of this gene product within the brain indicates that it may be involved in neuronal survival; synapse formation; conductance; neural differentiation, etc. Such involvement may impact many processes, such as learning and cognition. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 557 of SEQ ID NO:53, b is an integer of 15 to 571, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

The gene encoding the disclosed cDNA is thought to reside on chromosome 9. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 9.

This gene is expressed primarily in embryo and fetal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the embryo and fetal tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. fetal tissues, developmental, cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in embryonic and fetal tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of developmental disorders. Furthermore, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that

this protein may play a role in the regulation of cellular division. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as,
5 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the abovelisted tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of the present
10 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1233 of SEQ ID NO:54, b is an integer of 15
15 to 1247, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

20

The gene encoding the disclosed cDNA is thought to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 2.

This gene is expressed primarily in infant brain, placenta, some immune tissues,
25 and, to a lesser extent, in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and immune disorders. Similarly, polypeptides and
30 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the early developmental stage tissues and immune tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues)
35 or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to

the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:204 as residues: Val-32 to Met-39, Leu-44 to Val-49.

- 5 The tissue distribution in fetal and immune tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of developmental and immune disorders. Furthermore, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within
- 10 embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene may be useful in the treatment of lymphoproliferative
- 15 disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. The protein product of this
- 20 gene is useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms,
- 25 hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in
- 30 synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a
- 35 tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more

- 5 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 834 of SEQ ID NO:55, b is an integer of 15 to 848, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

- When tested against Jurkat T-cells, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates T-cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.
- 15
20

This gene is expressed primarily in brain tissues, and to a lesser extent, in T-cells.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neuronal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, brain, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 25
30
35

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:205 as residues: Ser-33 to Ser-44.

The tissue distribution in T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neuronal and immune system disorders. Furthermore, expression of this gene product in T-cells, as well as the observed biological activity of this gene product, indicates that this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the expression within brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 655 of SEQ ID NO:56, b is an integer of 15 to 669, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

When tested against U937 Myeloid cell lines, supernatants removed from cells
10 containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. When tested against K562 leukemia cell lines, supernatants removed from cells containing this gene activated the ISRE assay. Thus, it
15 is likely that this gene activates leukemia cells through the Jak-STAT signal transduction pathway. The interferon-sensitive response element is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the
20 binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. Contact of cells with supernatant expressing the product of this gene increases the permeability of bovine chondrocytes to calcium. Thus, it is likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product of this gene binds a receptor on the surface of
25 the chondrocyte cells. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating bone cells.

This gene is expressed primarily in breast and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
30 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pregnancy disorders including miscarriage. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast and placenta, expression of this
35 gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. placental tissues, breast, bone, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another

tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 The tissue distribution in both placenta and breast indicates a role for this protein in the treatment and/or detection of miscarriages in suspect individuals, of birth defects, of breast cancer, and female infertility. Furthermore, the biological assay data strongly indicates that the translation product of this gene is actively involved in the initiation of several signal transduction pathways and the activation of several cell types.

10 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more
15 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 666 of SEQ ID NO:57, b is an integer of 15 to 680, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

The gene encoding the disclosed cDNA is thought to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage
25 analysis for chromosome 11. One embodiment of this gene comprises the polypeptides of the following amino acid sequence:

MWGQPRPVDSVWSSSIPKKSVESNDNKSHLHKREH (SEQ ID NO:362),

MTTKAIFTKGNIDSLSFKSNMWSVYI (SEQ ID NO:363). An additional

embodiment is the polynucleotides encoding these polypeptides.

30 This gene is expressed primarily in the pancreas.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pancreatic related disorders such as diabetes. Similarly, polypeptides and
35 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this

gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. pancreas, endocrine, metabolic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, bile, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene in pancreatic tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment/detection of endocrine disorders and metabolic disorders associated with the pancreas including diabetes, pancreatitis, and pancreatic cancer. Furthermore, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothalamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 510 of SEQ ID NO:58, b is an integer of 15 to 524, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

This gene is expressed primarily in chondrosarcoma tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, including diseases of the skeletal system, particularly with respect to the

cartilagenous structures and also cancer of these tissues. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. bone, connective, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in chondrosarcoma tumors indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment/diagnosis of cartilage disorders including arthritis and cancer. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 413 of SEQ ID NO:59, b is an integer of 15 to 427, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

The translation product of this gene shares sequence homology with sorbin which is thought to be important in the manufacture of vitamin C. Additionally, sorbin is thought to be important in the process of stimulating water and electrolyte absorption in various cells in the body. Porcine Sorbin has activity in stimulating water and electrolyte absorption across mucosa. It has been pursued as a regulator of electrolyte absorption in the nasal and enteric mucosa. This gene was identified in hypothalamus suggesting that it could play a role in the CNS regulation of water or electrolyte absorption.

This gene is expressed primarily in human hypothalamus tissue from a patient suffering from Alzheimer's disease.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurologic disorders (eg. Alzheimer's disease). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:209 as residues: Leu-29 to Leu-37, Gln-65 to Asp-70, Gln-85 to Gly-95.

The tissue distribution in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of Alzheimer's disease. Additionally, the translation product of this gene, based upon its homology to the porcine sorbin, could be useful for the detection and/or amelioration of disorders involving the CNS regulation of water or electrolyte absorption. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1249 of SEQ ID NO:60, b is an integer of 15 to 1263, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 51

This gene is expressed primarily in synovium, and to a lesser extent, in other tissues.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, synovial diseases such as synovial sarcoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes
10 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the synovium, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. connective tissues, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample
5 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in synovium indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of
10 synovial diseases such as arthritis. Furthermore, the expression of this gene product in synovium would suggest a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g. trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders
5 such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
0 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the
5 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 706 of SEQ ID NO:61, b is an integer of 15 to 720, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:61, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

This gene is expressed primarily in immune tissues and fast-growing tissues, such as tumor and early-stage developmental tissues, and, to a lesser extent, in some other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and growth related disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune tissues and fast-growing tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. developmental, immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:211 as residues: Ala-28 to Ala-47.

The tissue distribution in immune tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of immune and growth related disorders. Furthermore, expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 575 of SEQ ID NO:62, b is an integer of 15 to 589, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

One embodiment of this gene comprises polypeptides of the following amino acid sequence: DSXLDRRPSGPDVKFLSNKHHFSMVC (SEQ ID NO:364). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in spleen, and to a lesser extent, in a range of hematopoietic cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. spleen, immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:212 as residues: Cys-25 to Trp-30.

The tissue distribution of this gene in spleen tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment or detection of immune or hematopoietic disorders including arthritis, asthma, immunodeficiency

diseases and leukemia. Expression of this gene product in spleen indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 672 of SEQ ID NO:63, b is an integer of 15 to 686, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

This gene is expressed primarily in human normal breast, and to a lesser extent, in dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, glandular problems involving cells of epithelial origin including breast cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell

type(s). For a number of disorders of the above tissues or cells, particularly of the female endocrine system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. breast, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:213 as residues: Ser-32 to Asn-44.

The tissue distribution in breast tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis or treatment of both malignant and non-malignant problems of the breast tissues, including cancer. Alternatively, the expression in dendritic tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 438 of SEQ ID NO:64, b is an integer of 15 to 452, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

When tested against U937 Myeloid cell lines, supernatants removed from cells
5 containing this gene activated the GAS assay. Thus, it is likely that this gene activates
myeloid cells through the Jak-STAT signal transduction pathway. The gamma
activating sequence (GAS) is a promoter element found upstream of many genes which
are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal
transduction pathway involved in the differentiation and proliferation of cells.
10 Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS
element, can be used to indicate proteins involved in the proliferation and differentiation
of cells.

This gene is expressed primarily in early stage human tissues, immune tissues,
and to a lesser extent, in other tissues such as prostate.

15 Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, development and immune related diseases. Similarly, polypeptides and
antibodies directed to these polypeptides are useful in providing immunological probes
20 for differential identification of the tissue(s) or cell type(s). For a number of disorders
of the above tissues or cells, particularly of the immune and early stage human tissues,
expression of this gene at significantly higher or lower levels may be detected in certain
tissues or cell types (e.g. immune, developmental, cancerous and wounded tissues) or
bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and
25 spinal fluid) or another tissue or cell sample taken from an individual having such a
disorder, relative to the standard gene expression level, i.e., the expression level in
healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in embryonic and immune tissues indicates that
polynucleotides and polypeptides corresponding to this gene are useful for diagnosis
30 and treatment of developmental and immune related diseases. The biological activity
data supports the assertion that the translation product of this gene is useful in the
treatment and/or diagnosis of diseases related to the immune system. Protein, as well
as, antibodies directed against the protein may show utility as a tissue-specific marker
and/or immunotherapy target for the above listed tissues.

35 Many polynucleotide sequences, such as EST sequences, are publicly available
and accessible through sequence databases. Some of these sequences are related to SEQ
ID NO:65 and may have been publicly available prior to conception of the present

invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 356 of SEQ ID NO:65, b is an integer of 15 to 370, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where b is greater than or equal to a + 14.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 56

The translation product of this gene shares sequence homology with medicago sativa salt-inducible protein.

This gene is expressed primarily in human chronic synovitis.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal or rheumatoid disorders, particularly, chronic synovitis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
20 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. connective tissues, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal
25 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:215 as residues: Lys-30 to Ser-44, Pro-77 to His-82.

30 The tissue distribution in synovium indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis of and as a therapeutic agent for chronic synovitis. In addition, the expression of this gene product in synovium would suggest a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders
35 afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as

dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 973 of SEQ ID NO:66, b is an integer of 15 to 987, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:66, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

The translation product of this gene shares high sequence homology with the rat and mouse peroxisomal membrane proteins [gil297437], which appears to play a crucial role in transporting proteins into the organelle. Some human genetic disorders involving peroxisome biogenesis, such as Zellweger syndrome, may be caused by genetic defects of the import machinery located in the peroxisomal membrane. When tested against fibroblast cell lines, supernatants removed from cells containing this gene activated the EGR1 assay. Thus, it is likely that this gene activates fibroblast cells through a signal transduction pathway. Early growth response 1 (EGR1) is a promoter associated with certain genes that induces various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

This gene is expressed primarily in normal human liver.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the hepatic system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic disorders and liver metabolic

diseases, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. liver, cancerous and wounded tissues) or bodily fluids (e.g. lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:216 as residues: Lys-57 to Ser-66.

The tissue distribution in liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of diseases relating to the liver. Furthermore, the homology indicates that the translational product of this gene may be useful in the detection and treatment of a number of disorders resulting from the improper transport of proteins into the organelle due to defects in peroxisomal membrane proteins, such as Zellweger syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1004 of SEQ ID NO:67, b is an integer of 15 to 1018, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

The gene encoding the disclosed cDNA is thought to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

This gene is expressed primarily in human fetal dura mater.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, developmental or neurologic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. brain, developmental, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:217 as residues: Ala-19 to Lys-34.

The tissue distribution in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurological diseases. Furthermore, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, and/or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 748 of SEQ ID NO:68, b is an integer of 15 to 762, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

The gene encoding the disclosed cDNA is thought to reside on chromosome 16.
5 Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 16.

This gene is expressed primarily in T helper cell and human uterine cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, relating to hemopoietic and uterus disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and female reproductive
15 system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level
20 in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-helper cells and uterine tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders relating to both the immune and female reproductive systems. Expression of this gene product in T-cells indicates a role in the
25 regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin,
30 the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product
35 may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell

types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 616 of SEQ ID NO:69, b is an integer of 15 to 630, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 60

This gene is expressed primarily in human fetal epithelium, and to a lesser extent, in testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental or reproductive disorders, in addition to diseases of the integumentary system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the diseases relating to the epithelium, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. epithelium, testes, developmental, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fetal epithelium and testes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of epithelium related diseases. In addition, polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles,

freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. Moreover, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athletes foot, and ringworm). Furthermore, the tissue distribution also indicates that the protein product of this gene is useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 926 of SEQ ID NO:70, b is an integer of 15 to 940, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where b is greater than or equal to a + 14.

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

When tested against both U937 Myeloid cell and Jurkat T-cell cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates both T-cells and myeloid cells through the Jak-STAT
5 signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins
10 involved in the proliferation and differentiation of cells.

This gene is expressed primarily in human adult lymph node and in early stage human lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
15 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, lymphatitis and pulmonary disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system
20 and respiratory system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue
25 or bodily fluid from an individual not having the disorder.

The tissue distribution in adult lymph indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of diseases relating to the immune system and respiratory system. Furthermore, expression of this gene product in lymph nodes indicates a role in the regulation of the
30 proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein,
35 as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency

diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The biological activity data supports the notion that the translation product of this gene is an activator of various cells of the immune system, and thus could play an important role in the activities of the immune system.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1089 of SEQ ID NO:71, b is an integer of 15 to 1103, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:71, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

This gene is expressed primarily in glioblastoma and anergic T-cell.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural and immune disorders, such as glioblastosis cerebri. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, neural, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of disorders relating to the CNS and the immune system. Furthermore, expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 885 of SEQ ID NO:72, b is an integer of 15 to 899, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:72, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

One embodiment of this gene comprises polypeptides of the following amino acid sequence:

35 CLAEAVSVIQSIPIFNETGRFSFTLPYPVKIKVRFSFFLQIYLIMIFLGLYINFRHLY
KQRRRRYGQKKKRSTKKKDLGFLPV (SEQ ID NO:365). An additional
embodiment is the polynucleotides encoding these polypeptides.